



comprising same and uses for diagnosing and preventing tuberculosis  
 Patent: WO 9909186-A 907 25-FEB-1999;  
 PORTNOI DENIS (FR); GUIGUENO AGNES (FR)

JOURNAL source  
 FEATURES Location/Qualifiers

1..1068  
 /organism="Mycobacterium tuberculosis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1773"  
 CDS  
 1..1068  
 /note="50D"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC05178.1"  
 /db\_xref="GI:9928796"  
 /translation="MSNSRRSLRWSSLLVLAAYGLGLATAPAAAPALSDQFAD  
 FPALPLDPSAMVAQPVQVNTKLGYNNAGCTGVIDPBGVLFNNHVIAGATD  
 INAFVSGSGQTVGVYDRTQDVAVLQLRAGGLPSAAAGGVAVGEPVAVMNSG  
 GGGTPRAVGRVAVLQTVQASDLSLTGAETLNLQIDAAIQPDGGGPPVAVNLGQ  
 VVGMTAAADNFQSGGGGFAIPIGQAMATAGTIRSGGSPVHIGTFAFLGLGVD  
 NNGGARVQRVVGSAASLGISTGDIVTAVDGPINSATAMADALNGHHPGDVISVT  
 WQTKSGGTGTGNVTLAEGPPA"

BASE COUNT 169 a 352 c 375 g 172 t  
 ORIGIN

Query Match 100.0%; Score 396; DB 6; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-57;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGCAGGATTCGCCATTCGGATC 60  
 DB 670 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGCAGGATTCGCCATTCGGATC 729  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCTGGGGGTTCACCCACCGTTTAT 120  
 DB 730 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCTGGGGGTTCACCCACCGTTTAT 789  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 180  
 DB 790 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 849  
 QY 181 GTCCAAACGGTGTGGGAGCGCTCCGGGCGCAAGTCTCGGATCTCCACCGCGGCGACG 240  
 DB 850 GTCCAAACGGTGTGGGAGCGCTCCGGGCGCAAGTCTCGGATCTCCACCGCGGCGACG 909  
 QY 241 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATCGGCGACGCTTAAC 300  
 DB 910 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATCGGCGACGCTTAAC 969  
 QY 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACG 360  
 DB 970 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACG 1029  
 QY 361 ACAGGGAAGTGACATTTGGCCGAGGACCCCGGCC 396  
 DB 1030 ACAGGGAAGTGACATTTGGCCGAGGACCCCGGCC 1065

RESULT 2  
 AX005790  
 LOCUS 1143 bp DNA linear PAT 24-AUG-2000  
 AX005790 Sequence 909 from Patent WO9909186.  
 AX005790  
 VERSION  
 AX005790.1 GI:9928797  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mycobacterium tuberculosis  
 Mycobacterium tuberculosis  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
 tuberculosis complex.

REFERENCE  
 AUTHORS  
 TITLE  
 1  
 Portnoi, D. and Guigueno, A.  
 Polypeptide nucleic sequences exported from mycobacteria, vectors  
 comprising same and uses for diagnosing and preventing tuberculosis

Patent: WO 9909186-A 909 25-FEB-1999;  
 PORTNOI DENIS (FR); GUIGUENO AGNES (FR)  
 Location/Qualifiers

1..1143  
 /organism="Mycobacterium tuberculosis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1773"  
 CDS  
 1..1143  
 /note="SEQ ID NO 50F"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC05179.1"  
 /db\_xref="GI:9928798"  
 /translation="KNPAARTLKAGTISDYPDTGGYMSNRSLRWSSLLSVLAA  
 VGLGLATAPAAAPALSDQFADFPALPLDPSAMVAQPVQVNTKLGYNNAGCTG  
 VIDPBGVLFNNHVIAGATDINAFVSGSGQTVGVYDRTQDVAVLQLRAGGLPS  
 AAAGGVAVGEPVAVMNSGCGGCTPRAPRGRVAVLQTVQASDLSLTGAETLNLQ  
 IDAAIQPDGGGPPVAVNLGQVVGMTAAADNFQSGGGGFAIPIGQAMATAGTIR  
 SGGSPVHIGTFAFLGLGVDNNGGARVQRVVGSAASLGISTGDIVTAVDGA  
 PINSATAMADALNGHHPGDVISVTWQTKSGGTGTGNVTLAEGPPA"

BASE COUNT 189 a 373 c 395 g 186 t  
 ORIGIN

Query Match 100.0%; Score 396; DB 6; Length 1143;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-57;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGCAGGATTCGCCATTCGGATC 60  
 DB 745 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGCAGGATTCGCCATTCGGATC 804  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCTGGGGGTTCACCCACCGTTTAT 120  
 DB 805 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCTGGGGGTTCACCCACCGTTTAT 864  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 180  
 DB 865 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 924  
 QY 181 GTCCAAACGGTGTGGGAGCGCTCCGGGCGCAAGTCTCGGATCTCCACCGCGGCGACG 240  
 DB 925 GTCCAAACGGTGTGGGAGCGCTCCGGGCGCAAGTCTCGGATCTCCACCGCGGCGACG 984  
 QY 241 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATCGGCGACGCTTAAC 300  
 DB 985 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATCGGCGACGCTTAAC 1044  
 QY 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACG 360  
 DB 1045 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACG 1104  
 QY 361 ACAGGGAAGTGACATTTGGCCGAGGACCCCGGCC 396  
 DB 1105 ACAGGGAAGTGACATTTGGCCGAGGACCCCGGCC 1140

RESULT 3  
 AR303127  
 LOCUS 2287 bp DNA linear PAT 12-JUN-2003  
 AR303127 Sequence 1 from patent US 6544522.  
 AR303127  
 ACCESSION  
 AR303127.1 GI:31691855  
 VERSION  
 AR303127.1  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Unknown.  
 Unclassified.  
 1 (bases 1 to 2287)  
 SKEIKY, F., Anderson, M. and Campos-Neto, A.  
 Fusion proteins of mycobacterium tuberculosis antigens and their  
 uses  
 Patent: US 6544522-A 1 08-APR-2003;  
 Location/Qualifiers

1..2287  
 JOURNAL source

BASE COUNT. 372 a 717 c 814 g 381 t 3 others

ORIGIN /organism="unknown"

Query Match 100.0%; Score 396; DB 6; Length 2287;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-57;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 60  
 DB 63 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 122  
 QY 61 GGGCAGCGATGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTCAT 120  
 DB 123 GGGCAGCGATGCGATCGCGGCCAGATCCGATCGGGTGGGGTTCACCCACCGTTCAT 182  
 QY 121 ATCGGGCTACCGCTTCCTCGGCTTGGGTGGTGTGTGACACAAACGCGCGCACGA 180  
 DB 183 ATCGGGCTACCGCTTCCTCGGCTTGGGTGGTGTGTGACACAAACGCGCGCACGA 242  
 QY 181 GTCCACGCGTGTGCGGAGCGCTCGCGCGGCAAGTCTCGGCATCTCCACCGCGGCGTG 240  
 DB 243 GTCCACGCGTGTGCGGAGCGCTCGCGCGGCAAGTCTCGGCATCTCCACCGCGGCGTG 302  
 QY 241 ATCACCGCGTGCAGCGCGCTCGGATCACTCGGCCACCGCGATGCGGCGGCTTAAC 300  
 DB 303 ATCACCGCGTGCAGCGCGCTCGGATCACTCGGCCACCGCGATGCGGCGGCTTAAC 362  
 QY 301 GGGCATCATCCCGTACGTCATCTCGGTGACCTGCGCAACCAACGCGCGCGCGT 360  
 DB 363 GGGCATCATCCCGTACGTCATCTCGGTGACCTGCGCAACCAACGCGCGCGCGT 422  
 QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
 DB 423 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 458

RESULT 4  
 MTCI418B 11700 bp DNA linear BCT 02-SEP-2002  
 LOCUS Mycobacterium tuberculosis H37Rv complete genome; segment 7/162.  
 DEFINITION 296071 ALI23456  
 ACCESSION 296071.1 GI:3242254  
 VERSION  
 KEYWORDS  
 SOURCE Mycobacterium tuberculosis H37Rv  
 ORGANISM Mycobacterium tuberculosis H37Rv  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
 tuberculosis complex.

REFERENCE 1  
 Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,  
 Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E.,  
 Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,  
 Connor, R., Davies, K., Devlin, K., Feltwell, T., Gentles, S.,  
 Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,  
 Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,  
 Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,  
 Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and  
 Barrall, B.G.  
 Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence  
 Nature 393 (6685), 537-544 (1998)

JOURNAL 98295987  
 MEDLINE  
 PUBMED 9634230  
 REFERENCE 2 (bases 1 to 11700)  
 AUTHORS Parkhill, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
 tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome  
 Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
 Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
 On Jun 20, 1998 this sequence version replaced gi:2181960.

Notes:  
 Details of M. tuberculosis sequencing at the Sanger Centre are  
 available on the World Wide Web.  
 (URL, <http://www.sanger.ac.uk/Projects/M.tuberculosis/>) CDS have  
 been renumbered from the original cosmid submissions but the old  
 gene designations are in brackets after the new gene numbers.  
 Gene prediction was based on a Hidden Markov Model of TB genes  
 implemented in TBparse (Krogh) supplemented with visual inspection  
 of positional base preference in codons, especially where there is  
 an increase in the observed/expected third position G + C.  
 CAUTION: In some cases we may not have predicted the correct  
 initiation codon. Where possible we choose an initiation codon  
 (atg, gtg, or ttg) which is preceded by an upstream ribosome  
 binding site sequence (optimally 5-13bp before the initiation  
 codon). If this cannot be identified we choose the most upstream  
 initiation codon.

FEATURES  
 source  
 Location/Qualifiers  
 1. .11700  
 /organism="Mycobacterium tuberculosis H37Rv"  
 /mol\_type="genomic DNA"  
 /strain="H37Rv"  
 /db\_xref="taxon:83332"  
 1309. .11293  
 /organism="Mycobacterium tuberculosis H37Rv"  
 /mol\_type="genomic DNA"  
 /strain="H37Rv"  
 /db\_xref="taxon:83332"  
 /clone="I418B"  
 10774. .>11700  
 /organism="Mycobacterium tuberculosis H37Rv"  
 /mol\_type="genomic DNA"  
 /strain="H37Rv"  
 /db\_xref="taxon:83332"  
 /clone="CI5"  
 misc\_feature  
 <1. .1409  
 /note="fragment designated v031. Does not represent a  
 physical clone"  
 83. .88  
 /note="possible RBS, AAGGAG, for Rv0119"  
 97. .1674  
 /gene="fadD7"  
 97. .1674  
 /gene="fadD7"  
 /note="Rv0119, (MTV031.13-MTCI418B.01), len: 525.fadD7,  
 PossibleCoenzyme A-ligase similar to 4-coumarate:CoA  
 ligase of many organisms e.g. PTU39405\_1 (537 aa). Contains  
 PS00455Putative AMP-binding domain signature. FASTA  
 scores: gp|U39405|PTU39405\_1 Pinus taeda xyiem  
 4-coumarate:CoA (537 aa).opt:483z-score: 526.1 E():  
 8.3e-22; 28.2% identity in 440 aaoverlap score is 0.896"  
 /codon\_start=1  
 /transl\_table=11  
 /product="fadD7"  
 /protein\_id="CAB09455.1"  
 /db\_xref="GI:3242255"  
 /db\_xref="SPTREMBL:O07169"  
 /translation="MASDFCPRIADLVEAATRLPEAPALVVTADRIAISHRDLRLV  
 DELAGOLTRSGLLPGDRVALRMGSNAEEFVALLAASRADLVVPLDLPALPITEORVRS  
 QARAGRVILIDAGPHDRAEPTTRWHPLVNVGSDGSPGGTSLVHLDATENPAT  
 TPLGLRPDDAMIMFTGTTGLPKMVPFWTHANIASVRAITGTRLSFRDTAVAMPLY  
 HGHLIASLATIASGGVSLPARGVSAHFDWDIKAVGATWYATVTHQTLERS  
 ATEPSGRPKAALRFIRCSAPLTAQAALQTEFAAPVAVCAFGMTATHTVTTQIEG  
 IDOTETPVYSTGLVGRSTGAQIRIVGSDGLPLPAGAVGEIWLGTWVRGYLGDPTIT  
 ANFTDGLWLTGDLGSLAAGDLSIRGKELINRGKESIPERVELASHPNVMEA  
 AVGVPHQLYGEAVAVIVPRESAPPTREELVQFCRLEAAFEIPASQASGLPHTA  
 KGSLLRRRAVRAEFGHSV"  
 610. .645  
 /gene="fadD7"  
 /note="PS00455 Putative AMP-binding domain signature,  
 [LIVMY] [21][STG][STAG][ST][STEI][SG].[PASLIYW][KR], info  
 count = 22.0"  
 complement(1675. .3819)  
 /gene="fusA2"

misc\_feature  
 610. .645  
 /gene="fadD7"  
 /note="PS00455 Putative AMP-binding domain signature,  
 [LIVMY] [21][STG][STAG][ST][STEI][SG].[PASLIYW][KR], info  
 count = 22.0"  
 complement(1675. .3819)  
 /gene="fusA2"





**TITLE** Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains

**JOURNAL** Unpublished

**REFERENCE** 2 (bases 1 to 14029)

**AUTHORS** Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.

**TITLE** Direct Submission

**JOURNAL** Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

**FEATURES** location/Qualifiers

1. .14029

/organism="Mycobacterium tuberculosis CDC1551"

/mol\_type="genomic DNA"

/strain="CDC1551"

/db\_xref="taxon:83331"

/note="clinical strain"

/complement(99..533)

/gene="MT0129"

/complement(99..533)

/gene="MT0129"

/note="identified by Glimmer2; putative"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

/protein\_id="AAK44353.1"

/db\_xref="GI:13879611"

/translation="MGEFDEKLRFAQSPVARLATSTPDGTHLVVVFALGARRPAEA TGAIVTAVDAKKTKQRLRLANLEHNPRASVLVDSYDDWTQLMWMRADGVAAIHRDGEVMRAAYLLRAKYTOYOSVPLNGPVIATIAVORWASRHA"

/complement(643..1062)

/gene="MT0130"

/complement(643..1062)

/gene="MT0130"

/note="identified by Glimmer2; putative"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

/protein\_id="AAK44354.1"

/db\_xref="GI:13879612"

/translation="MASSSLPLLRHCVSDDQTVTVVGFDDGLGKTARRIAALVQORA IFNDNRNTAVHSGDDAVLGHVLPVGPDPDPLHASSMYSLGMCOSVNGRPDDATA LVSVRCHVQTDPTDSCGGRDRPQLPCAPLDYHRHH"

/complement(1047..1415)

/gene="MT0131"

/complement(1047..1415)

/gene="MT0131"

/note="identified by match to protein family HMM"

/codon\_start=1

/transl\_table=11

/product="DNA-binding protein, CopG family"

/protein\_id="AAK44355.1"

/db\_xref="GI:13879613"

/translation="MTKKRNPADYVIGDDVEYSDVLKQEEVYVDGERITDERVEOM ASLSRLAREANLIPGKSLSGGSAHSPAVQVVSFKATHAKLKELRKMSVSKL LRVLDFEVRQRENTGRILPRR"

/complement(1682..3367)

/gene="MT0132"

/complement(1682..3367)

/gene="MT0132"

/note="similar to GB:AL123456; identified by sequence similarity; putative"

/codon\_start=1

/transl\_table=11

/product="PE\_PGRS family protein"

/protein\_id="AAK44356.1"

/db\_xref="GI:13879614"

/translation="MFAGGAGGLRCVMSFVSVAPEIVVAAATDLAGISATSAANA AAAAPTAVLAAGADEVSAIAALFSGHAQYALSAQAAAFHQQFVQTLAGGAGAYA AAQAQVEQOILLAINAPTQALLGRPLIGNAGADGAPGTQAGAGGGLYNGNGGSGA AGQAGGAGGAGLIGHGGSGAGSGAAGGGHGLWNGGVGGSGAGAGVAG

GHGAGGAAGLWAGAGGGGNGGADANITVSGDGLGGAGGGGGLYVDGAGGGHGG QGAGGAGGAGGGGGGAGRLWGTGGAGGGGGTGPPLPGAGMAAGGAGG LINGGAGGGGCVGAGVGAGXGAGGNAMLIHGAGGAGGDSFANGAGGAGGAG GHLFGNGSGGHHGAVTAGNTGIGGAGVGGDARLIGHGAGGAGGRAGALVGRDGG PGNGGAGGOLYNGGDDGGCGGQAGFANNIGGTGGAGCGNGCPAILSGNGGCGAG GAGGAGGAGGAGGCVGAGGAGGAGGTGGTLOAAVSLVTLFALFAPGQPGDTGPGP"

3519..4586

/gene="MT0133"

3519..4586

/gene="MT0133"

/note="identified by match to protein family HMM"

/codon\_start=1

/transl\_table=11

/product="serine protease, putative"

/protein\_id="AAK44357.1"

/db\_xref="GI:13879615"

/translation="MSNRRLSLWSLLSLVLAAGVLGLATAPAAQAPALSDREFAD FPALPLPSAMVAQVQGVNINIKLGTNNVAGAGTGIVDPNGVLTNNHVIAGATD INAFVSGGOTYGVVDYDRTQDVQLGRAGGLPSAAIGGVAVGEPVVMNGSG GGGTTPRAGFVRVVALGTVQASDSLTAETLNLQFDAAIQPGSGGVPVWNGLG VVGMNTAASDNFOLSGGQGAIPIGQAMATAGQIRSGGSGPTVHIGPTAFGLGVVD NNGNGARVORVVGSAAPASLIGISTGDIVTAVDGAIPINSATAMADALNGHHPGDIVST WORKSGGTRTGNVTLAEPPA"

4695..6500

/gene="MT0134"

4695..6500

/gene="MT0134"

/note="similar to GP:2808807; identified by sequence similarity; putative"

/codon\_start=1

/transl\_table=11

/product="alpha-amylase family protein"

/protein\_id="AAK44358.1"

/db\_xref="GI:13879616"

/translation="MNEAEHSVEHPVQGSHEVGVVEHPDAKDFGSAALPADPTWF KHAFVYEVLRFFDASDAGSLRGLIDRLDYLQWLGIDCWLPPYDPSLRGGYD IRDFYKVLPEFGTVDDFVALVDAARRGRIITLDMNTHSESHPFQESRRDPDGPY GDYVWSDTSERYTDARIIFVDTEESNMSFDPVRQYWHRRFFSHQPLNDYDPAVQE AMIDVIRFWLGLGIDGFLDADVPYLFEREGTNCNLPETHAFKRVKRVDDPEPGVR LLAENQWPGDVVEYFGDPNTGDECHMAHFPLMPRIFMAVRRESFPISEIIAQTP PIPDMAQGIPLRNHDELTEMVTDDEERDYMAEYAKDPRKANVGIRRRRLAILLDND RNQIELTALLSLPGSPVLYGDEIGMDVWLGDRDGVRIPMQWTFDRNAGSTAN RQVLPSPQDPVYGVQAVNEAQRDSTSLNFTRLMVAVRHRPAPAVAGFQELGG SNPSVLAVRQVAGDGGTVLCVNNLSRFQPIELDLQOMNTYTPVELTGHVEFPRIG QVPYLLPLPGHGYWFQTLTTEVGAPPTCGGERL"

6603..7970

/gene="MT0135"

6603..7970

/gene="MT0135"

/note="similar to GP:2808801; identified by sequence similarity; putative"

/codon\_start=1

/transl\_table=11

/product="pep2 protein"

/protein\_id="AAK44359.1"

/db\_xref="GI:13879617"

/translation="MTRSDTLATKLPSDWMLPRQWYAGNRRELATVKPGYVVALRHN LDVLVDVTVTDGATYQVVLGWDFPASEYGTAAIGVADRTGDALYDVAGPQF LLSLIVSSAVCGTSTGEVTFTRPDELPPAAQPRVCDAEQSTSVIFDRRAILKVPFR RVSSGINPDIELNRLVTRAGNPHVRLILGAYQFCGRPNRSPDALAYALGVYEAANA AEGWAMATSVRDLFAEGDLFAHEVGGDFAGESRGLSEAVASVATLADSLGTQAATF PVDRLARLSSTVAVVPELREYAPTIEQFQKLAEEAITVORVHGDHLHQLGVLRPES WLLDIFEGEPQPLDERRAPDPLRVAGVLSRFYAAAYGPLVDQADKQLAARARE VERRAAFCDGAYAVASGIDPRDSALLLGVELDKAVYETGYETRRHPCWLPILURISA RLTAAS"

8038..8817

/gene="MT0136"

8038..8817

/gene="MT0136"

/note="identified by Glimmer2; putative"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

```
/protein_id="AAK44360.1"
/db_xref="GI:13879618"
/translation="MOREIYDGEARLSWLAALAGILGATAFTHSAGYFVFTMGNSQ
RAVLGFGDAMSVTASLILFFVAGVVIASVCRHFWAHPGPTVLTTLTSLIFAA
GVDIMLGWHEHSMDFPILFVFGIGALNTSPVKDGEVSPVLSYVLYKMGQIE
RHLAGKVEDWLGCFYLLHASFVLCGAAGGAISNVVTGPQNLAAVVAACATGTYTLH
ADRGLVNRKPPGPKRLPRALRGELDSGTSTPATNYGSS"
complement(8949..9989)
/gene="MT0137"
/complement(8949..9989)
/gene="MT0137"
/notes="similar to SP:P31953 CB:X57229 PID:48828
PID:1567734 PID:1877254; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="esterase, putative, antigen 85-C"
/protein_id="AAK44361.1"
/db_xref="GI:13879619"
/translation="MGLVYAMTFFEOVRRLRSAAATLPRRLAIAAMGAVLVYGLVGT
GGPATAGASRECLPVEYLQVPSASGRDVKVQGGGPHAVYLLDGLRAQDDYNGWD
INTPAFEYQSGLSVIMPVQSGSFYTDWQFOSGONHYTKWETFLREMPAWLQ
ANKGVPTGNAAYLSMGGSGSALILAAVYPQFPVAAASLGSFLNPSEGHWPTLGLAM
NDSGGVYNSMGPSPSDPAKRNPMVQIPRLVANNTRIWYCGNGTSGDLGDNIIPA
KELEGILTNFTNTRDYAADGRNGVFNPPNGTHSWPWNQELVAMKADIQHVING
ATPPAPAAPAA"
10218..10673
/gene="MT0138"
10218..10673
/gene="MT0138"
/notes="similar to SP:P08634 PID:46222; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="nodulation protein N-related protein"
/protein_id="AAK44362.1"
/db_xref="GI:13879620"
```

```
Query Match
Best Local Similarity 100.0%; Score 396; DB 1; Length 14029;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCGGATACCTCCAGCTGTCCTCCAGGTGGCGAGGATTCGCCATTCGGATC 60
Db |||||||
QY 4188 ACGGCGGCTCGGATACCTCCAGCTGTCCTCCAGGTGGCGAGGATTCGCCATTCGGATC 4247
Db |||||||
QY 61 GGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTGCACACCGCGTTCAT 120
Db |||||||
QY 4248 GGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTGCACACCGCGTTCAT 4307
Db |||||||
QY 121 ATCGGCGCTACCGCTTCCTCGGCTTGGTGTGTCGACACACGCGCACGACGA 180
Db |||||||
QY 4308 ATCGGCGCTACCGCTTCCTCGGCTTGGTGTGTCGACACACGCGCACGACGA 4367
Db |||||||
QY 181 GTCCACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGGACGTG 240
Db |||||||
QY 4368 GTCCACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGGACGTG 4427
Db |||||||
QY 241 ATCACCGCGGTGAGCGGCTCCGATCACTCGGCACCGCGATCGGCGACGCGCTTAAC 300
Db |||||||
QY 4428 ATCACCGCGGTGAGCGGCTCCGATCACTCGGCACCGCGATCGGCGACGCGCTTAAC 4487
Db |||||||
QY 301 GGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGACGCGT 360
Db |||||||
QY 4488 GGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGACGCGT 4547
Db |||||||
QY 361 ACAGGAGCGTCACATTCGCCGAGGACCGCCCGGCC 396
Db |||||||
QY 4548 ACAGGAGCGTCACATTCGCCGAGGACCGCCCGGCC 4583
Db |||||||
```

RESULT 6  
BX248334

```
LOCUS
DEFINITION
BX248334
Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment
1/14.
ACCESSION
BX248334
VERSION
BX248334.1
KEYWORDS
complete genome.
SOURCE
Mycobacterium bovis subsp. bovis AF2122/97
Mycobacterium bovis subsp. bovis AF2122/97
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Garnier,T., Eiglmeier,K., Camus,J.-C., Medina,N., Mansoor,H.,
Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsemp,C., Simon,S.,
Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L.,
Wheeler,P.R., Parkhill,J., Barrell,B.G., Cole,S.T., Gordon,S.V. and
Hewinson,G.
The complete genome sequence of Mycobacterium bovis
Online Publication
PNAS 10.1073/pnas.1130426100 ( Microbiology )
2 (bases 1 to 343050)
Garnier,T.
Direct Submission
Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex
15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the
Mycobacterium bovis sequencing teams, TB Research Group, Veterinary
Laboratories Agency Weybridge, Woodham Lane, New Haw, Addlestone,
Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus,
Hinxton,Cambridge CB10 1SA, UK. P74 Annotation, Genopole, Institut
Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France.
Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28
rue du Docteur Roux, 75724 Paris Cedex 15, France
Location/Qualifiers
FEATURES
Source
1..343050
/organism="Mycobacterium bovis subsp. bovis AF2122/97"
/mol_type="genomic DNA"
/strain="AF2122/97"
/db_xref="taxon:233413"
1..1524
/gene="dnaA"
/locus_tag="Mb0001"
1..1524
/gene="dnaA"
/locus_tag="Mb0001"
/notes="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001,
len: 507 aa, from Mycobacterium tuberculosis strain H37Rv,
(99.6% identity in 507 aa overlap). dnaA, chromosomal
replication initiator protein (see citations below),
equivalent to other Mycobacterial CHROMOSOMAL REPLICATION
INITIATOR PROTEINS e.g. P46388|DnaA_MYCLE from
Mycobacterium leprae (502 aa); Q9L7L7|DnaA_MYCPA from
Mycobacterium paratuberculosis (509 aa); P49990|DnaA_MYCAV
from Mycobacterium avium (508 aa); P49992|DnaA_MYCSM from
Mycobacterium smegmatis (504 aa); etc. Also highly similar
to others except in N-terminus e.g. Q9ZH75|DnaA_STRCH
CHROMOSOMAL REPLICATION INITIATOR PROTEIN from
Streptomyces chrysomallus (624 aa); Q9ZH76|DnaA_STRRE from
Streptomyces reticuli (643 aa); DnaA_ECOLI|P03004|B3702
chromosomal replication initiator protein from Escherichia
coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0,
(43.2% identity in 389 aa overlap); etc. Contains P500017
ATP/GTP-binding site motif A (P-loop) and P501008 DnaA
protein signature. BELONGS TO THE DnaA FAMILY. Note that
the first base of this gene has been taken as base 1 of
the Mycobacterium bovis genomic sequence."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DnaA"
/protein_id="CAD92863.1"
/db_xref="GI:3161673"
/translation="MTDDPGSGFTTWNNAVVELNGDPKVDGPPSSDANLSAPLTPQQ
```



Db 152188 GTCCAAACGGCTGGTGGGAGCGCTCGGGGGCAAGTCTCGGCATCTCCACCGGCGACGGT 152247  
QY 241 ATCACCGCGTGCAGCGCGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 300  
Db 152248 ATCACCGCGTGCAGCGCGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 152307  
QY 301 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGT 360  
Db 152308 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGT 152367  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 396  
Db 152368 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 152403

RESULT 7  
AR169152  
LOCUS AR169152 447 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 4 from patent US 6290969.  
ACCESSION AR169152  
VERSION AR169152.1 GI:17906927  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S. and Twardzik, D.R.  
TITLE Compounds and methods for immunotherapy and diagnosis of  
tuberculosis  
JOURNAL Patent: US 6290969-A 4 18-SEP-2001;  
FEATURES Location/Qualifiers  
source 1. 447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;  
Best Local Similarity 99.7%; Pred. No. 9.2e-57;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 60  
Db 11 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTTCACCCACCGCTTAC 120  
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTTCACCCACCGCTTAC 130  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGGCGACGA 180  
Db 131 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGGCGACGA 190  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
QY 241 ATCACCGCGTGCAGCGCGCTCCGATCACTACCGCCACCGCGATCGCGGCGGACGGT 300  
Db 251 ATCACCGCGTGCAGCGCGCTCCGATCACTACCGCCACCGCGATCGCGGCGGACGGT 310  
QY 301 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGT 360  
Db 311 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGT 370  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 396  
Db 371 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 406

RESULT 8  
AR182442  
LOCUS AR182442 447 bp DNA linear PAT 20-APR-2002

DEFINITION Sequence 4 from patent US 6338852.  
ACCESSION AR182442  
VERSION AR182442.1 GI:20225649  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S. and Twardzik, D.R.  
TITLE Compounds and methods for diagnosis of tuberculosis  
JOURNAL Patent: US 6338852-A 4 15-JAN-2002;  
FEATURES Location/Qualifiers  
source 1. 447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;  
Best Local Similarity 99.7%; Pred. No. 9.2e-57;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 60  
Db 11 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTTCACCCACCGCTTAC 120  
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTTCACCCACCGCTTAC 130  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGGCGACGA 180  
Db 131 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGGCGACGA 190  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
QY 241 ATCACCGCGTGCAGCGCGCTCCGATCACTACCGCCACCGCGATCGCGGCGGACGGT 300  
Db 251 ATCACCGCGTGCAGCGCGCTCCGATCACTACCGCCACCGCGATCGCGGCGGACGGT 310  
QY 301 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGT 360  
Db 311 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGT 370  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 396  
Db 371 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 406

RESULT 9  
AR194825  
LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 4 from patent US 6350456.  
ACCESSION AR194825  
VERSION AR194825.1 GI:20244262  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed, S.G., Skeiky, Y.A.W. and Dillon, D.C.  
TITLE Compositions and methods for the prevention and treatment of M.  
tuberculosis infection  
JOURNAL Patent: US 6350456-A 4 26-FEB-2002;  
FEATURES Location/Qualifiers  
source 1. 447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;

```
Best Local Similarity 99.7%; Pred. No. 9.2e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCCGCTCCGATTAATTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 60
Db 11 ACGGCCGCTCCGATTAATTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 70
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 120
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 130
Qy 121 ATCGGGCTACCGCTTCCTCGCTTGGGTGGTGTGTCACAAACGCGACGGCGCAGCA 180
Db 131 ATCGGGCTACCGCTTCCTCGCTTGGGTGGTGTGTCACAAACGCGACGGCGCAGCA 190
Qy 181 GTCCAAACGCTGGTTCGGGAGCGGTCCCGGGGCAAGTCTCGCATCTCCACGGGAGCTG 240
Db 191 GTCCAAACGCTGGTTCGGGAGCGGTCCCGGGGCAAGTCTCGCATCTCCACGGGAGCTG 250
Qy 241 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 300
Db 251 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 310
Qy 301 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 360
Db 311 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 370
Qy 361 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 396
Db 371 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 406

RESULT 10
AR233097
LOCUS AR233097 447 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 4 from patent US 6458366.
ACCESSION AR233097
VERSION AR233097.1 GI:27275533
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skelky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
TITLE Compounds and methods for diagnosis of tuberculosis
JOURNAL Patent: US 6458366-A 4 01-OCT-2002;
FEATURES Location/Qualifiers
source
1..447
/organism="unknown"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9.2e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCCGCTCCGATTAATTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 60
Db 11 ACGGCCGCTCCGATTAATTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 70
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 120
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 130
Qy 121 ATCGGGCTACCGCTTCCTCGCTTGGGTGGTGTGTCACAAACGCGACGGCGCAGCA 180
Db 131 ATCGGGCTACCGCTTCCTCGCTTGGGTGGTGTGTCACAAACGCGACGGCGCAGCA 190
Qy 181 GTCCAAACGCTGGTTCGGGAGCGGTCCCGGGGCAAGTCTCGCATCTCCACGGGAGCTG 240
Db 191 GTCCAAACGCTGGTTCGGGAGCGGTCCCGGGGCAAGTCTCGCATCTCCACGGGAGCTG 250
Qy 241 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 300
Db 251 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 310
Qy 301 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 360
Db 311 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 370
Qy 361 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 396
Db 371 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 406

RESULT 12
BD006325
```

```
Qy 241 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 300
Db 251 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 310
Qy 301 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 360
Db 311 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 370
Qy 361 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 396
Db 371 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 406

RESULT 11
AX429596
LOCUS AX429596 447 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 4 from Patent EP1203817.
ACCESSION AX429596
VERSION AX429596.1 GI:21540845
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Reed,S.G., Skelky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,
Vedvick,T.S. and Twardzik,D.R.
TITLE Compounds and methods for immunotherapy and diagnosis of
tuberculosis
JOURNAL Patent: EP 1203817-A 4 08-MAY-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source
1..447
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9.2e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCCGCTCCGATTAATTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 60
Db 11 ACGGCCGCTCCGATTAATTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 70
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 120
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 130
Qy 121 ATCGGGCTACCGCTTCCTCGCTTGGGTGGTGTGTCGACAAACGCGACGGCGCAGCA 180
Db 131 ATCGGGCTACCGCTTCCTCGCTTGGGTGGTGTGTCGACAAACGCGACGGCGCAGCA 190
Qy 181 GTCCAAACGCTGGTTCGGGAGCGGTCCCGGGGCAAGTCTCGCATCTCCACGGGAGCTG 240
Db 191 GTCCAAACGCTGGTTCGGGAGCGGTCCCGGGGCAAGTCTCGCATCTCCACGGGAGCTG 250
Qy 241 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 300
Db 251 ATCACCGCGTTCGACGGCGTTCGATCAATTCGGCCACCGCATGGCGGAGCGGCTTAAC 310
Qy 301 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 360
Db 311 GGGCATCATCCCGTACGTCATCTCGGTGACCTGACCTGGCAACCAACGCGACCGCT 370
Qy 361 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 396
Db 371 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 406

RESULT 12
BD006325
```

LOCUS	BD006325	447 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Compounds and methods for diagnosis of Tuberculosis.				
ACCESSION	BD006325				
VERSION	BD006325.1	GI:18634696			
KEYWORDS	JP 2001500383-A/4.				
SOURCE	unidentified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 447)				
AUTHORS	Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R., Vedvick, T.S., Twardzik, D.R. and Iodes, M.J.				
TITLE	Compounds and methods for diagnosis of Tuberculosis				
JOURNAL	Patent: JP 2001500383-A 4 16-JAN-2001;				
COMMENT	CORIXA CORP				
	OS unidentified				
	PN JP 2001500383-A/4				
	PD 16-JAN-2001				
	PF 07-OCT-1997 JP 1998518432				
	PR 11-OCT-1996 US 08/729622, 13-MAR-1997 US 08/818111 PI				
	STEVEN G REDD, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS				
	NETO, PI				
	RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI				
	MICHAEL J IODES				
	PC C12N15/31.C07K14/35.C07K16/12.C12Q1/68.C12N15/62.G01N33/53 CC				
	Strandedness: Single;				
	CC Topology: Linear;				
	EH Key Location/Qualifiers				
	FT source 1..447				
	FT Location/Qualifiers				
	1..447				
	/organism="Unidentified".				
FEATURES					
source					
	1..447				
	/organism="Unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
BASE COUNT	79 a 146 c 149 g 72 t 1 others				
ORIGIN					
Query Match	99.6%;	Score 394.4;	DB 6;	Length 447;	
Best Local Similarity	99.7%;	Pred. No. 9.2e-57;			
Matches	395;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	ACGGCCGCGTCGATTAACCTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60			
Db	11	ACGGCCGCGTCGATTAACCTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 70			
QY	61	GGCAGCGCATGCGGATCGGGCCAGATCCGATCGGTGGGGGGTACCCACCGTTCAT 120			
Db	71	GGCAGCGCATGCGGATCGGGCCAGATCCGATCGGTGGGGGGTACCCACCGTTCAT 130			
QY	121	ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGCAACGGCGCACGA 180			
Db	131	ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGCAACGGCGCACGA 190			
QY	181	GTCCAAACGGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240			
Db	191	GTCCAAACGGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250			
QY	241	ATCACCGGCTGACCGGCTCCGATCAACTCGGCCACCGCGCATGCGGACGCGCTTAAC 300			
Db	251	ATCACCGGCTGACCGGCTCCGATCAACTCGGCCACCGCGCATGCGGACGCGCTTAAC 310			
QY	301	GGGCATCATCCGGTCACTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACCGGT 360			
Db	311	GGGCATCATCCGGTCACTCATCTCGGTGAACTGGCAACCAAGTCGGCGGCGACCGGT 370			
QY	361	ACAGGAACGTGACATTTGCCGAGGACCCCGGCC 396			
Db	371	ACAGGAACGTGACATTTGCCGAGGACCCCGGCC 406			
RESULT 13					
BD006445					
LOCUS	BD006445	447 bp	DNA	linear	PAT 31-JAN-2002

DEFINITION	Compounds and methods for immunotherapy and diagnosis of Tuberculosis.			
ACCESSION	BD006445			
VERSION	BD006445.1	GI:18634816		
KEYWORDS	JP 2001501832-A/4.			
SOURCE	unidentified#			
ORGANISM	unclassified#			
REFERENCE	1. (bases 1 to 447)			
AUTHORS	Reed S.G., Skeiky Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R., Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.			
TITLE	Compounds and methods for immunotherapy and diagnosis of			
JOURNAL	Patent: JP 2001501832-A 4 13-FEB-2001;			
COMMENT	CORIXA CORP			
	OS Unidentified#			
	PN JP 2001501832-A/4			
	PD 13-FEB-2001			
	PF 07-OCT-1997 JP 1998518456			
	PR 11-OCT-1996 US 08/730510,13-MAR-1997 US 08/918112 PI			
	STEVEN G REED, YASIR A W SKEIKY,DAVIN C DILLON, PI ANTONIO CAMPOS			
	NETO,			
	PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI			
	MICHAEL J LODES			
	PC C12N15/31.C07K14/35.A61K39/04.A61K48/00.A61K49/00.C12N15/62,			
	PC C07K19/00,			
	PC G01N33/50.G01N33/60.G01N33/569.C12N1/19.C12N1/20.C12N1/21, PC			
	C12N5/10//			
	PC (C12N1/21.C12N1:19)			
	CC Strandedness: Single;			
	CC Topology: Linear;			
	FH Key	Location/Qualifiers		
	FT source	1..447		
	FT	Location/Qualifiers	/organism='Unidentified'.	
FEATURES	source	1..447		
		/organism='unidentified'		
		/mol_type='genomic DNA'		
		/db_xref='taxon:32644'		
BASE COUNT	79 a 146 c	149 g	72 t	1 others
ORIGIN				
	Query Match	99.6%;	Score 394.4;	DB 6;
	Best Local Similarity	99.7%;	Pred.No.9.2e-57;	
	Matches 399;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCGCATC 60		
Db	11	ACGGCCGCGTCCGATAAATTTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCGCATC 70		
Qy	61	GGCAGCGGATCCGGATCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 120		
Db	71	GGCAGCGGATCCGGATCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 130		
Qy	121	ATCGGGCTACGCCCTTCTCGGCTTGGGTGTGTCTGCACAACAAACGGCAACGGCGCACGA 180		
Db	131	ATCGGGCTACGCCCTTCTCGGCTTGGGTGTGTCTGCACAACAAACGGCAACGGCGCACGA 190		
Qy	181	GTCCAAACGGGTGTCTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 240		
Db	191	GTCCAAACGGGTGTCTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250		
Qy	241	ATCACCGCGGTGCACCGCGCTCCGATCACTCGGCCACCGCGATGGCGGCGGCTTAAC 300		
Db	251	ATCACCGCGGTGCACCGCGCTCCGATCACTCGGCCACCGCGATGGCGGCGGCTTAAC 310		
Qy	301	GGGCATCATCCGGGTGACGTCACTCTCGGTGACCTGGCAAAACCAAGTCTGGCGCGCACGCGT 360		
Db	311	GGGCATCATCCGGGTGACGTCACTCTCGGTGACCTGGCAAAACCAAGTCTGGCGCGCACGCGT 370		
Qy	361	ACAGGAAAGTGACATTTGCCCGAGGACCCCGGCC 396		
Db	371	ACAGGAAAGTGACATTTGCCCGAGGACCCCGGCC 406		

RESULT 13  
BD006445  
LOCUS

```

RESULT 14
BD069285
LOCUS
DEFINITION
  BD069285
  Compounds and methods for immunotherapy and diagnosis of
  tuberculosis.
ACCESSION
  BD069285
VERSION
  JP 2001517069-A/4.
KEYWORDS
  unidentified
SOURCE
  unidentified
ORGANISM
  unclassified.
REFERENCE
  1 (bases 1 to 447)
  Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
  Vedvick,T.H. and Twardzik,D.R.
  Compounds and methods for immunotherapy and diagnosis of
  tuberculosis.
  Patent: JP 2001517069-A 4 02-OCT-2001;
  JOURNAL
  CORIXA CORP
  OS Unidentified
  PI NETO,
  PN JP 2001517069-A/4
  PD 02-OCT-2001
  PR 30-AUG-1996 JP 1997511464
  PR 01-SEP-1995 US 08/523436,22-SEP-1995 US 08/533634 PR
  22-MAR-1996 US 08/620874,05-JUN-1996 US 08/659683 PR
  12-JUL-1996 US 08/680574
  PI STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
  PI NETO,
  PI RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
  C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
  C12N5/10,
  PC C12N1/21/A61K39/04,(C12N1/21,C12R1:19)
  CC Strandedness: Single;
  CC Topology: Linear;
  CC Compounds and methods for immunotherapy and diagnosis of CC
  tuberculosis
  FH Key
  FT source
  FT Location/Qualifiers
  1..447
  /organism="Unidentified".
  /organism="unidentified"
  /mol_type="genomic DNA"
  /db_xref="taxon:32644"
  79 a 146 c 149 g 72 t 1 others
  BASE COUNT
  ORIGIN
  Query Match 99.68; Score 394.4; DB 6; Length 447;
  Best Local Similarity 99.7%; Pred. No. 9.2e-57;
  Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 60
  Db 11 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 70
  Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 120
  Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 130
  Qy 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCACACACACGCGGACGCGACGA 180
  Db 131 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCACACACACGCGGACGCGACGA 190
  Qy 181 GTCAACCGCTGGTTCGGAGCGGTCCCGCGGCAAGTCTCGGCATCTCCACCGGACGCTG 240
  Db 191 GTCAACCGCTGGTTCGGAGCGGTCCCGCGGCAAGTCTCGGCATCTCCACCGGACGCTG 250
  Qy 241 ATCAACCGCTGGTTCGGAGCGGTTCGGATCAACTCGGCCACCGCATTCGGGCGGCGCTTAA 300
  Db 251 ATCAACCGCTGGTTCGGAGCGGTTCGGATCAACTCGGCCACCGCATTCGGGCGGCGCTTAA 310
  Qy 301 GGGCATCATCCCGTACGCTCATCTCGGTACCTTGGGCAACCAACCAAGTCCGGGCGGACGCT 360
  Db 311 GGGCATCATCCCGTACGCTCATCTCGGTACCTTGGGCAACCAACCAAGTCCGGGCGGACGCT 370

```

```

Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
Db 371 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 406

RESULT 15
AR169165
LOCUS
DEFINITION
  AR169165
  Sequence 17 from patent US 6290969.
ACCESSION
  AR169165
VERSION
  AR169165.1 GI:17906947
KEYWORDS
  SOURCE
  Unknown.
ORGANISM
  Unclassified.
REFERENCE
  1 (bases 1 to 1872)
  Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
  Vedvick,T.S. and Twardzik,D.R.
  Compounds and methods for immunotherapy and diagnosis of
  tuberculosis
  Patent: US 6290969-A 17 18-SEP-2001;
  JOURNAL
  Location/Qualifiers
  FEATURES
  source
  1..1872
  /organism="unknown"
  BASE COUNT 318 a 616 c 604 g 332 t 2 others
  ORIGIN
  Query Match 98.8%; Score 391.2; DB 6; Length 1872;
  Best Local Similarity 99.2%; Pred. No. 2.4e-56;
  Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
  Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 60
  Db 758 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 817
  Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 120
  Db 818 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTAT 877
  Qy 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCACACACACGCGGACGCGACGA 180
  Db 878 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCACACACACGCGGACGCGACGA 937
  Qy 181 GTCAACCGCTGGTTCGGAGCGGTCCCGCGGCAAGTCTCGGCATCTCCACCGGACGCTG 240
  Db 938 GTCAACCGCTGGTTCGGAGCGGTCCCGCGGCAAGTCTCGGCATCTCCACCGGACGCTG 997
  Qy 241 ATCAACCGCTGGTTCGGAGCGGTTCGGATCAACTCGGCCACCGCATTCGGGCGGCGCTTAA 300
  Db 998 ATCAACCGCTGGTTCGGAGCGGTTCGGATCAACTCGGCCACCGCATTCGGGCGGCGCTTAA 1057
  Qy 301 GGGCATCATCCCGTACGCTCATCTCGGTACCTTGGGCAACCAACCAAGTCCGGGCGGACGCT 360
  Db 1058 GGGCATCATCCCGTACGCTCATCTCGGTACCTTGGGCAACCAACCAAGTCCGGGCGGACGCT 1117
  Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
  Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

Search completed: September 5, 2003, 07:52:35
Job time : 1680 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:00 ; Search time 246 Seconds  
(without alignments)  
4345.437 Million cell updates/sec

Title: US-09-684-215A-3

Perfect score: 396

Sequence: 1 acggcgctcgataactt.....tgccgagggaccccgcc 396

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	396	AA140769	Nucleotide sequenc
2	396	100.0	672	AA140772	Nucleotide sequenc
3	396	100.0	702	AA220206	Mycobacterium tube
4	396	100.0	702	AA140770	Nucleotide sequenc
5	396	100.0	702	ABK14140	DNA encoding antig
6	396	100.0	1002	AA140777	Mycobacterium tube
7	396	100.0	1002	AA140778	Mycobacterium tube
8	396	100.0	1002	AA228336	Mycobacterium sp.

9	396	100.0	1002	24	AA228337	Mycobacterium spec
10	396	100.0	1068	20	AA344251	Mycobacterium spec
11	396	100.0	1143	20	AA344252	Mycobacterium spec
12	396	100.0	1742	23	AA140771	Nucleotide sequenc
13	396	100.0	2190	24	AA140784	Mycobacterium sp.
14	396	100.0	2190	24	AA228343	Mycobacterium spec
15	396	100.0	2191	23	AA140773	Nucleotide sequenc
16	396	100.0	2286	24	ABK14128	DNA encoding antig
17	396	100.0	2287	20	AA220194	Mycobacterium tube
18	396	100.0	2287	20	AA140783	Mycobacterium spec
19	396	100.0	2287	24	AA228342	Mycobacterium spec
20	396	100.0	2808	24	AA147110	Mycobacterium sp.
21	396	100.0	4403765	22	AA199683	Mycobacterium tube
22	396	100.0	4411529	22	AA199682	Mycobacterium tube
23	394.4	99.6	447	18	AA191466	Mycobacterium tube
24	394.4	99.6	447	18	AA191403	Mycobacterium tube
25	394.4	99.6	447	19	AA191450	M. tuberculosis im
26	394.4	99.6	447	19	AA191450	M. tuberculosis im
27	394.4	99.6	447	20	AA191452	M. tuberculosis an
28	394.4	99.6	447	20	AA191452	M. tuberculosis re
29	394.4	99.6	447	22	AA191452	M. tuberculosis DN
30	394.4	99.6	447	24	AA191452	Mycobacterium tube
31	394.4	99.6	447	24	AA191452	Mycobacterium spec
32	394.4	99.6	1629	22	AA191452	M. tuberculosis DN
33	392.8	99.2	1871	23	AA191452	Nucleotide sequenc
34	391.2	98.8	1872	18	AA191477	Mycobacterium tube
35	391.2	98.8	1872	18	AA191477	Mycobacterium tube
36	391.2	98.8	1872	19	AA191477	M. tuberculosis im
37	391.2	98.8	1872	19	AA191477	Mycobacterium tube
38	391.2	98.8	1872	20	AA191477	M. tuberculosis an
39	391.2	98.8	1872	20	AA191477	M. tuberculosis re
40	391.2	98.8	1872	22	AA191477	M. tuberculosis DN
41	391.2	98.8	1872	24	AA191477	Mycobacterium tube
42	391.2	98.8	1872	24	AA191477	Mycobacterium tube
43	384	97.0	1067	22	AA191477	Mycobacterium spec
44	354	89.4	675	22	AA191477	Human /M. tubercul
45	354	89.4	675	22	AA191477	Ral2-P510S-C const

#### ALIGNMENTS

#### RESULT 1

AA140769

ID AA140769 standard; DNA; 396 BP.

XX AA140769;

AC AA140769;

XX 03-OCT-2002 (first entry)

XX Nucleotide sequence encoding Ral2 protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;

XX vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;

XX immunogen; cytokine; gene; ds.

XX Unidentified.

OS Key

XX Location/Qualifiers

PH CDS

FT 1..396

FT /\*tag= a

FT /product= "Ral2 protein"

FT /note= "No start or stop codon"

XX WO200125401-A2.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27652.

XX 07-OCT-1999; 99US-0158585.

XX (CORI-) CORIXA CORP.

```
XX PI Skeiky Y, Guderian J;
XX
XX WPI; 2001-266299/27.
XX P-PSDB; AAO22138.
XX
XX Recombinant nucleic acid molecule for producing high yield expression
PT of desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide
PT
XX
XX Claim 1; Fig 2; 39pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ral2 protein.
XX
XX Sequence 396 BP; 58 A; 132 C; 134 G; 52 T; 0 other;
XX
XX Query Match 100.0%; Score 396; DB 23; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 60
Db 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 60
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
Db 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCAGCA 180
Db 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCAGCA 180
QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
Db 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
QY 241 ATCACCGGGTGCAGCGGGCTCCGATCACTCGGCACCGCGATGGCGGACGCGTTTAA 300
Db 241 ATCACCGGGTGCAGCGGGCTCCGATCACTCGGCACCGCGATGGCGGACGCGTTTAA 300
QY 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGGT 360
Db 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGGT 360
QY 361 ACAGGGAAGGTGACATTTGGCGAGGACCCCGGCC 396
Db 361 ACAGGGAAGGTGACATTTGGCGAGGACCCCGGCC 396
XX
XX RESULT 2
XX AAL40772
XX ID AAL40772 standard; DNA; 672 BP.
XX
XX AC AAL40772;
XX
XX DT 03-OCT-2002 (first entry)
XX
```

```
DE Nucleotide sequence encoding Ral2-mammaglobin fusion protein.
XX
XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
XX Chimeric - Mammalian.
OS Chimeric - Unidentified.
XX
XX Key Location/Qualifiers
FT 4..666
FT /*tag= a
FT /product= "Ral2-mammaglobin fusion protein"
XX
XX WO200125401-A2.
XX
XX 12-APR-2001.
XX
XX 06-OCT-2000; 200WO-US27652.
XX
XX 07-OCT-1999; 99US-0158585.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Guderian J;
XX
XX WPI; 2001-266299/27.
XX P-PSDB; AAO22141.
XX
XX Recombinant nucleic acid molecule for producing high yield expression
PT of desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide
PT
XX
XX Disclosure; Fig 5; 39pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ral2-mammaglobin fusion protein.
XX
XX Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 other;
XX
XX Query Match 100.0%; Score 396; DB 23; Length 672;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 60
Db 25 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 84
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
Db 85 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 144
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCAGCA 180
Db 145 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCAGCA 204
QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
XX
```

Db 205 GTCCAAACCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTTCGGCATCTCCACCGGCGACGTG 264  
QY 241 ATCACCGGGTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300  
Db 265 ATCACCGGGTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 324  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGGCAACCAAGTCGGGCGGCGACGCGT 360  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGGCAACCAAGTCGGGCGGCGACGCGT 384  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 396  
Db 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 420

## RESULT 3

AAZ20206  
ID AAZ20206 standard; DNA; 702 BP.

XX AAZ20206;

XX 17-JAN-2000 (first entry)

DE Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.  
XX Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD;  
KW diagnosis; therapy; vaccine; immunogen; ss.

XX Mycobacterium tuberculosis.

PN W09951748-A2.

XX 14-OCT-1999.

XX 07-APR-1999; 99WO-US07717.

PR 07-APR-1998; 98US-0056556.

PR 30-DEC-1998; 98US-0223040.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Alderson M, Campos-Neto A;

XX WPI; 1999-601610/51.

DR P-PSDB; AAY32071.

XX New fusion proteins useful for diagnosis, prevention and treatment of tuberculosis -  
XX Example; Fig 13A-B; 83pp; English.  
XX This DNA sequence includes a coding region for a recombinant  
CC Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),  
CC termed Mtb24, composed of the antigens Ra12 and DPPD. The  
CC DNA is useful for the recombinant production of the fusion protein.  
CC Coding sequences for the antigens were modified by PCR in order  
CC to facilitate their fusion and subsequent expression of the fusion  
CC protein, and then ligated. The invention provides fusion proteins  
CC (see AAY32059-71) containing at least 2 M. tuberculosis antigens. The  
CC new fusion proteins and polynucleotides encoding them are useful as  
CC vaccines for preventing tuberculosis (claimed), for diagnosis (via  
CC in vitro assays or intradermal skin tests for detection of anti-M.  
CC tuberculosis antibodies), monitoring of disease progression, and  
CC treatment of tuberculosis. They are more effective immunogens than  
CC mixtures of the individual protein components.

XX Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;  
XX Query Match 100.0%; Score 396; DB 20; Length 702;  
XX Best Local Similarity 100.0%; Pred. No. 1.2e-84;  
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGCGATC 60

Db 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGCGATC 84  
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGTCCACCGTTCAT 120  
Db 85 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGTCCACCGTTCAT 144  
QY 121 ATCGGGCTACCGCTTCTCTCGGCTTGGGTGTTGTCGACAAACAAACGGCAACGGCGACGA 180  
Db 145 ATCGGGCTACCGCTTCTCTCGGCTTGGGTGTTGTCGACAAACAAACGGCAACGGCGACGA 204  
QY 181 GTCCAAACGCTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 205 GTCCAAACGCTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264  
QY 241 ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300  
Db 265 ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 324  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACAAAGTCGGGCGGCGACGCGT 360  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACAAAGTCGGGCGGCGACGCGT 384  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 396  
Db 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 420

## RESULT 4

AAZ40770  
ID AAZ40770 standard; DNA; 702 BP.

XX AAZ40770;

XX 03-OCT-2002 (first entry)

DE Nucleotide sequence encoding Ra12-DPPD fusion protein.  
XX Ra12; serine protease antigen; Mtb32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.

XX Chimeric - Unidentified.

XX Key Location/Qualifiers

XX CDS 4..696

XX FT /\*tag= a

XX FT /product= "Ra12-DPPD fusion protein"

XX PN W0200125401-A2.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27652.

XX 07-OCT-1999; 99US-0158585.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Guderian J;

XX WPI; 2001-266299/27.

XX P-PSDB; AAO22139.

XX Recombinant nucleic acid molecule for producing high yield expression  
XX of desired fusion polypeptides, encodes fusion polypeptide comprising  
XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide  
XX Example 1; Fig 3; 39pp; English.  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
XX fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14

CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2-DPPD fusion protein.  
 XX  
 SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

Query Match 100.0%; Score 396; DB 23; Length 702;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-84;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGATC 60  
 DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGATC 84  
 QY 61 GGGCAGGCGATGGCGATCGCGGCGGCGAGATCCGATCGGTGGGCGTCAACCCACCGTTTCAT 120  
 DB 85 GGGCAGGCGATGGCGATCGCGGCGGCGAGATCCGATCGGTGGGCGTCAACCCACCGTTTCAT 144  
 QY 121 ATCCGCGCTACCGCTTCTCCGGCTGGGTGTTGTCGACAAACGCGGCGGCGACGA 180  
 DB 145 ATCCGCGCTACCGCTTCTCCGGCTGGGTGTTGTCGACAAACGCGGCGGCGACGA 204  
 QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGGCGGCAAGTCTCGGATCTCCACCGCGGCGACG 240  
 DB 205 GTCCAAACGGTGGTGGGAGCGCTCCGGGCGGCAAGTCTCGGATCTCCACCGCGGCGACG 264  
 QY 241 ATCCGCGGTGCGAGCGGCTCCGATCACTCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 DB 265 ATCCGCGGTGCGAGCGGCTCCGATCACTCGGCGGCGGCGGCGGCGGCGGCGGCGG 324  
 QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGGCGGCGGCGG 360  
 DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGGCGGCGGCGG 384  
 QY 361 ACAGGGAAGTGCATTTGGCGGAGGACCCCGCGCC 396  
 DB 385 ACAGGGAAGTGCATTTGGCGGAGGACCCCGCGCC 420

## RESULT 5

ABK14140  
 ID ABK14140 standard; DNA; 702 BP.

XX AC ABK14140;

XX DT 08-MAY-2002 (first entry)

XX DE DNA encoding antigenic fusion protein Ral2-DPPD (Mtb24).

XX KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;  
 XX KW tuberculostatic; immunogen; vaccine; Ral2-DPPD; Mtb24.

XX OS Chimeric - Mycobacterium tuberculosis.

XX FH Key Location/Qualifiers

FT CDS 1..702

FT FT /tag= a

FT FT /product= "Mtb24 #1"

FT FT /partial

FT FT /transl\_except= (pos:694..696, aa:Xaa)

FT FT /note= "No start or stop codon. Xaa= In frame stop codon"

FT CDS 2..702

FT FT /tag= b  
 FT FT /product= "Mtb24 #2"  
 FT FT /partial  
 FT FT /note= "No start or stop codon"  
 FT FT /transl\_except= (pos:263..265, aa:Xaa)  
 FT FT /transl\_except= (pos:353..355, aa:Xaa)  
 FT FT /transl\_except= (pos:395..397, aa:Xaa)  
 FT FT /transl\_except= (pos:470..472, aa:Xaa)  
 FT FT /transl\_except= (pos:701..702, aa:Ser)  
 FT FT /note= "This codon has an apparent 1 nucleotide deletion  
 FT FT which alters the reading frame."  
 FT FT Xaa= In frame stop codon"  
 FT CDS 3..701  
 FT FT /tag= c  
 FT FT /product= "Mtb24 #3"  
 FT FT /partial  
 FT FT /note= "No start or stop codon"  
 FT FT /transl\_except= (pos:1..2, aa:Pro)  
 FT FT /transl\_except= (pos:39..41, aa:Xaa)  
 FT FT /transl\_except= (pos:321..323, aa:Xaa)  
 FT FT /transl\_except= (pos:339..341, aa:Xaa)  
 FT FT /transl\_except= (pos:450..452, aa:Xaa)  
 FT FT /transl\_except= (pos:621..623, aa:Xaa)  
 FT FT /note= "No start or stop codon. Xaa= In frame stop codon"  
 XX

PN US2002009459-A1.

XX 24-JAN-2002.

XX 07-APR-1999; 99US-0287849.

XX 13-MAR-1997; 97US-0818112.

PR 01-OCT-1997; 97US-0942578.

PR 18-FEB-1998; 98US-0025197.

PR 07-APR-1998; 98US-0056556.

PR 30-DEC-1998; 98US-0223040.

XX (REED/) REED S G.

PA (SKEI/) SKEIKY Y A.

PA (DILL/) DILLON D C.

PA (ALDE/) ALDERSON M.

PA (CAMP/) CAMPOS-NETO A.

PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;

XX WPI; 2002-171134/22.

XX P-PSDB; AAU74600, AAU76541, AAU76542.

XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for  
 XX diagnosing, treating or preventing M. tuberculosis infection, .  
 XX particularly as vaccine for treating or preventing tuberculosis .

XX Examples: Fig 13; 62pp; English.

XX The invention relates to a purified polypeptide which induces an immune  
 XX response of Mycobacterium tuberculosis. Polypeptides of the invention are  
 XX useful for diagnosing, treating or preventing M. tuberculosis infection,  
 XX particularly tuberculosis infection. In particular, the polypeptides are  
 XX useful as a vaccine formulation with an adjuvant to afford long-term  
 XX protection in animals against the development of tuberculosis. The  
 XX protein coding sequence may be used to encode a protein product for use  
 XX as an immunogen to induce and/or enhance an immune response to M.  
 XX tuberculosis. This sequence represents DNA encoding an M. tuberculosis  
 XX fusion protein of the invention. This polynucleotide encodes 3 different  
 XX proteins, each in a different reading frame.

SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

Query Match 100.0%; Score 396; DB 24; Length 702;

Best Local Similarity 100.0%; Pred. No. 1.2e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGATC 60

```
Db 25 ACAGCGCGTCCGATAAATTCAGCTGTCCAGAGGTGGCAGGATTCCGCAATTCGATC 84
Qy 61 GGGCAGGCGATGGCGATCGCGGGCCAGATCGGATGGGTGGGGTCAACCGTTTCAT 120
Db 85 GGGCAGGCGATGGCGATCGCGGGCCAGATCGGATGGGTGGGGTCAACCGTTTCAT 144
Qy 121 ATCGGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGCACGA 180
Db 145 ATCGGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGCACGA 204
Qy 181 GTCCAAACGCTGGTTCGGAGGCGTCCGGCGCAAGTCTCGCATCTCCACGGCGAGTG 240
Db 205 GTCCAAACGCTGGTTCGGAGGCGTCCGGCGCAAGTCTCGCATCTCCACGGCGAGTG 264
Qy 241 ATCACCAGCTGCGACGGCGTCCGATCAACTCGGCCACCGGATGGGCGCGCTTAAC 300
Db 265 ATCACCAGCTGCGACGGCGTCCGATCAACTCGGCCACCGGATGGGCGCGCTTAAC 324
Qy 301 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGAGTGGGCGGCGACGGT 360
Db 325 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGAGTGGGCGGCGACGGT 384
Qy 361 ACAGGAACTGACATTCGGCGAGGACCGCCCGGCC 396
Db 385 ACAGGAACTGACATTCGGCGAGGACCGCCCGGCC 420
```

## RESULT 6

```
AAD47077
ID AAD47077 standard; DNA; 1002 BP.
XX
AC AAD47077;
XX
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen;
KW gene; ds.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT CDS 4..996
FT FT /*tag= a
FT FT /product= "Ra35 mature antigenic protein"
FT FT /transl_except= (pos:547..549, aa:Asp)
FT FT /transl_except= (pos:550..552, aa:Ser)
```

## WO200272792-A2.

```
XX
XX AAD47078
XX ID AAD47078 standard; DNA; 1002 BP.
XX AC AAD47078;
XX
XX DT 27-JAN-2003 (first entry)
XX
XX DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
XX
XX KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
XX KW gene; antigen; ds.
XX
XX OS Mycobacterium tuberculosis.
XX
XX FH Key Location/Qualifiers
XX FT CDS 4..996
XX FT FT /*tag= a
XX FT FT /product= "Ra35FLMutSA mutant antigenic protein"
XX
XX PN WO200272792-A2.
XX
XX PD 19-SEP-2002.
XX
XX PF 13-MAR-2002; 2002WO-US08223.
XX
XX PR 13-MAR-2001; 2001US-275837P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Skeiky Y, Brannon M, Guderian J;
XX
XX DR WPI; 2002-759844/82.
XX
XX DR P-PSDB; AAE29702.
XX
XX PT New recombinant nucleic acid molecule comprising a Leishmania TSA,
XX PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective
XX PT immunity against pathogenic microorganisms e.g. Leishmania and
XX PT Mycobacterium tuberculosis
XX
XX PS Disclosure; Page 79; 155pp; English.
XX
XX CC The invention relates to a recombinant nucleic acid molecule encoding a
```

```
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides,
CC as in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is
CC M. tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen
CC encoding DNA.
XX
```

Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;

```
Query Match 100.0%; Score 396; DB 24; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGCGCGTCCGATAAATTCAGCTGTCCAGAGGTGGCAGGATTCCGCAATTCGATC 60
Db 598 ACAGCGCGTCCGATAAATTCAGCTGTCCAGAGGTGGCAGGATTCCGCAATTCGATC 657
Qy 61 GGGCAGGCGATGGCGATCGCGGGCCAGATCGGATGGGTGGGGTCAACCGTTTCAT 120
Db 658 GGGCAGGCGATGGCGATCGCGGGCCAGATCGGATGGGTGGGGTCAACCGTTTCAT 717
Qy 121 ATCGGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGCACGA 180
Db 718 ATCGGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGCGCACGA 777
Qy 181 GTCCAAACGCTGGTTCGGAGGCGTCCGGCGCAAGTCTCGCATCTCCACGGCGAGTG 240
Db 778 GTCCAAACGCTGGTTCGGAGGCGTCCGGCGCAAGTCTCGCATCTCCACGGCGAGTG 837
Qy 241 ATCACCAGCTGCGACGGCGTCCGATCAACTCGGCCACCGGATGGGCGCGCTTAAC 300
Db 838 ATCACCAGCTGCGACGGCGTCCGATCAACTCGGCCACCGGATGGGCGCGCTTAAC 897
Qy 301 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGAGTGGGCGGCGACGGT 360
Db 898 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGAGTGGGCGGCGACGGT 957
Qy 361 ACAGGAACTGACATTCGGCGAGGACCGCCCGGCC 396
Db 958 ACAGGAACTGACATTCGGCGAGGACCGCCCGGCC 993
```

## RESULT 7

```
AAD47078
ID AAD47078 standard; DNA; 1002 BP.
XX
XX AC AAD47078;
XX
XX DT 27-JAN-2003 (first entry)
XX
XX DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
XX
XX KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
XX KW gene; antigen; ds.
XX
XX OS Mycobacterium tuberculosis.
XX
XX FH Key Location/Qualifiers
XX FT CDS 4..996
XX FT FT /*tag= a
XX FT FT /product= "Ra35FLMutSA mutant antigenic protein"
XX
XX PN WO200272792-A2.
XX
XX PD 19-SEP-2002.
```

```

XX PF 13-MAR-2002; 2002WO-US08223.
XX PR 13-MAR-2001; 2001US-275837P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky Y, Brannon M, Guderian J;
XX DR WPI: 2002-759844/82.
XX DR P-PSDB; AAE29703.
XX PT New recombinant nucleic acid molecule comprising a Leishmania TSA,
XX PT Leif, M15 or 6H polynucleotide, useful as vaccine to elicit protective
XX PT immunity against pathogenic microorganisms e.g. Leishmania and
XX PT Mycobacterium tuberculosis
XX PS Disclosure; Page 80-81; 155pp; English.
XX CC The invention relates to a recombinant nucleic acid molecule encoding a
XX CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
XX CC polynucleotide sequence encoding an antigen or an antigenic fragment from
XX CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
XX CC polypeptide or its fragment. The Leishmania polynucleotide is selected
XX CC from TSA, Leif, M15, and 6H polynucleotides. Sequences of the invention
XX CC are used in methods for eliciting immune response in mammals. They are
XX CC useful as vaccines to elicit protective immunity against pathogenic
XX CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
XX CC polypeptides are used for enhancing the expression of polynucleotides,
XX CC as in vivo diagnostic agents and for raising antibodies in a non-human
XX CC animal. The invention is used in gene therapy. The present sequence is
XX CC M. tuberculosis Ra35FLmutSA mutant antigenic protein encoding DNA.
XX SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;

Query Match 100.0%; Score 396; DB 24; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGGCGGCTCGGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 60
DB 598 ACGGGCGGCTCGGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 657
QY 61 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCA 120
DB 658 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCA 717
QY 121 ATCGGGCTTACCGCTTCTCTCGGCTTGGGTGTCGACAAACGCGGACGCGACGA 180
DB 718 ATCGGGCTTACCGCTTCTCTCGGCTTGGGTGTCGACAAACGCGGACGCGACGA 777
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGGGCAAGTCTGGGATCTCCACCGGCGACGTG 240
DB 778 GTCCAAACGGTGGTGGGAGCGCTCCGGGGCAAGTCTGGGATCTCCACCGGCGACGTG 837
QY 241 ATCACCGGCTGGACGGCGCTCCGATCACTCGGCCACCGCATGGCGACGCGCTTAA 300
DB 838 ATCACCGGCTGGACGGCGCTCCGATCACTCGGCCACCGCATGGCGACGCGCTTAA 897
QY 301 GGGCATCATCCGGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGGGCGGACGCG 360
DB 898 GGGCATCATCCGGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGGGCGGACGCG 957
QY 361 ACAGGAGCGTGACATTTGGCGGAGGACCCCGGCC 996
DB 958 ACAGGAGCGTGACATTTGGCGGAGGACCCCGGCC 993

RESULT 8
AAD28336
ID AAD28336 standard; cDNA; 1002 BP.
XX AC AAD28336;

```

```

XX DT 22-APR-2002 (first entry)
XX DE Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX KW Fusion protein; antigen; serological sensitivity; immune response;
XX KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX OS Mycobacterium sp.
XX FH Key Location/Qualifiers
XX FT CDS 4..996
XX FT /*tag= a
XX FT /product= "Ra35 mature protein"
XX FT /transl_except= (pos:547..549, aa:Asp)
XX FT /transl_except= (pos:550..552, aa:Ser)
XX PN WO200198460-A2.
XX PD 27-DEC-2001.
XX PF 20-JUN-2001; 2001WO-US19959.
XX PR 20-JUN-2000; 2000US-0597796.
XX PR 01-FEB-2001; 2001US-265737P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky Y, Reed S, Alderson M;
XX DR WPI: 2002-147798/19.
XX DR P-PSDB; AAE17566.
XX PT Composition comprising MTB39 antigen and MTB32A antigen from
XX PT Mycobacterium species, useful for eliciting immune response in a
XX PT subject
XX PS Disclosure; Page 95; 136pp; English.
XX CC The present invention relates to fusion proteins containing at least
XX CC two Mycobacterium species antigens, nucleotides encoding them and
XX CC compositions comprising such fusion proteins. The present invention
XX CC particularly relates to nucleic acids encoding fusion proteins that
XX CC include two or more individual M. tuberculosis antigens which increase
XX CC the serological sensitivity of sera from individuals infected with
XX CC tuberculosis and methods for their use in diagnosis, prevention and
XX CC treatment of tuberculosis infection. Sequences of the invention are
XX CC useful for eliciting an immune response in a mammal, e.g., human,
XX CC immunised with BCG. They are useful in the diagnosis, treatment and
XX CC prevention of Mycobacterium infection. The fusion proteins and the
XX CC polynucleotides are useful as diagnostic tools in patients infected
XX CC with Mycobacterium, in vitro and in vivo assays for detecting humoral
XX CC antibodies or cell-mediated immunity against M. tuberculosis, for the
XX CC diagnosis of an infection or monitoring of disease progression, as
XX CC immunogens to generate or elicit a protective immune response in a
XX CC patient and for raising anti-M. tuberculosis antibodies in a non-human
XX CC animal. Sequences of the invention are also used as vaccines. MTB32A
XX CC fusion proteins of the invention are useful as in vivo diagnostic agents
XX CC for intradermal skin test. The present sequence is a cDNA encoding
XX CC Mycobacterium species MTB32A (Ra32FL) mature protein.
XX SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;

Query Match 100.0%; Score 396; DB 24; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGGCGGCTCGGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 60
DB 598 ACGGGCGGCTCGGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 657
QY 61 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCA 120

```

Db 658 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTACCCACCGCTTCAT 717  
 Qy 121 ATCGGGGCTACCGCTTCCTCGGCTTGGGTGTGACAAACGCGAACGCGGACGA 180  
 Db 718 ATCGGGGCTACCGCTTCCTCGGCTTGGGTGTGACAAACGCGAACGCGGACGA 777  
 Qy 181 GTCAACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 240  
 Db 778 GTCAACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 837  
 Qy 241 ATCACCGCGTGTGCGGAGCGCTCGGCATCACTCGCCACCGCGATGGGCGCGCTTAAC 300  
 Db 838 ATCACCGCGTGTGCGGAGCGCTCGGCATCACTCGCCACCGCGATGGGCGCGCTTAAC 897  
 Qy 301 GGGCATCATCCCGTGTGCGGAGCGCTCGGCATCACTCGCCACCGCGATGGGCGCGCTTAAC 360  
 Db 898 GGGCATCATCCCGTGTGCGGAGCGCTCGGCATCACTCGCCACCGCGATGGGCGCGCTTAAC 957  
 Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
 Db 958 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 993

## RESULT 9

AAD28337

ID AAD28337 standard; cDNA; 1002 BP.

XX

AC AAD28337;

XX

DT 22-APR-2002 (first entry)

DE

Myobacterium species Ra35FLMutSA mutant cDNA.

XX

Fusion protein; antigen; serological sensitivity; immune response;  
tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.

XX

OS Myobacterium sp.

XX

FH Key Location/Qualifiers

FT CDS 4..996

FT FT /\*tag= a

FT FT /product= "Ra35FLMutSA protein"

XX

PN W0200198460-A2.

XX

PD 27-DEC-2001.

XX

XX 20-JUN-2001; 2001WO-US19959.

XX

PR 20-JUN-2000; 2000US-0597796.

PR

PR 01-FEB-2001; 2001US-265737P.

XX

XX (CORI-) CORIXA CORP.

XX

PI Skeiky Y, Reed S, Alderson M;

XX

XX WPI; 2002-147798/19.

DR

DR P-PSDB; AAEL17567.

XX

XX Composition comprising MTB39 antigen and MTB32A antigen from

PT

Myobacterium species, useful for eliciting immune response in a

PT

subject -

XX

XX Claim 67; Page 96-97; 136pp; English.

PS

XX The present invention relates to fusion proteins containing at least  
two Myobacterium species antigens, nucleotides encoding them and  
compositions comprising such fusion proteins. The present invention  
particularly relates to nucleic acids encoding fusion proteins that  
include two or more individual M. tuberculosis antigens which increase  
the serological sensitivity of sera from individuals infected with  
tuberculosis and methods for their use in diagnosis, prevention and  
treatment of tuberculosis infection. Sequences of the invention are

CC

CC useful for eliciting an immune response in a mammal, e.g., human,  
 CC immunised with BCG. They are useful in the diagnosis, treatment and  
 CC prevention of Mycobacterium infection. The fusion proteins and the  
 CC polynucleotides are useful as diagnostic tools in patients infected  
 CC with Mycobacterium, in vitro and in vivo assays for detecting humoral  
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
 CC diagnosis of an infection or monitoring of disease progression, as  
 CC immunogens to generate or elicit a protective immune response in a  
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
 CC animal. Sequences of the invention are also used as vaccines. MTB32A  
 CC fusion proteins of the invention are useful as in vivo diagnostic agents  
 CC for intradermal skin test. The present sequence is a cDNA encoding  
 CC Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA.  
 XX

SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;

Query Match 100.0%; Score 396; DB 24; Length 1002;

Best Local Similarity 100.0%; Pred. NO. 1.2e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGCATC 60

Db 598 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGCATC 657

Qy 61 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTGACCCACCGCTTCAT 120

Db 658 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTGACCCACCGCTTCAT 717

Qy 121 ATCGGGCCTACCGCTTCTCGGCTTGGGTGTTGTCACAAACGCGAACGGCGCACGA 180

Db 718 ATCGGGCCTACCGCTTCTCGGCTTGGGTGTTGTCACAAACGCGAACGGCGCACGA 777

Qy 181 GTCAACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGCGAGCTG 240

Db 778 GTCAACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGCGAGCTG 837

Qy 241 ATCACCGCGTGTGCGGAGCGCTCGGCATCACTCGGCCACCGCGATGGGCGGACGGCGCTTAAC 300

Db 838 ATCACCGCGTGTGCGGAGCGCTCGGCATCACTCGGCCACCGCGATGGGCGGACGGCGCTTAAC 897

Qy 301 GGGCATCATCCCGTGTGCGGAGCGCTCGGCATCACTCGGCATCGCCACCGCGCTTAAC 360

Db 898 GGGCATCATCCCGTGTGCGGAGCGCTCGGCATCACTCGGCATCGCCACCGCGCTTAAC 957

Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396

Db 958 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 993

## RESULT 10

AA34251

ID AA34251 standard; DNA; 1068 BP.

XX

AC AA34251;

XX

DT 06-JUL-1999 (first entry)

XX

DE Mycobacterium species nucleic acid sequence 50D.

XX

Secreted protein; Mycobacterium; primer; PCR; amplification; probe;  
hybridisation; detection; vaccine; immunisation; infection; ss.

XX

OS Mycobacterium sp.

XX

XX WO9909186-A2.

XX

XX 25-FEB-1999.

XX

XX 14-AUG-1998; 98WO-FR01813.

XX

XX 11-SEP-1997; 97FR-0011325.

PR

PR 14-AUG-1997; 97FR-0010404.

XX

```

PA (INSP ) INST PASTEUR.
XX
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
XX
PR 11-SEP-1997; 97FR-0011325.
PR 14-AUG-1997; 97FR-0010404.
XX
PA (INSP ) INST PASTEUR.
XX
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
XX
DR WPI: 1999-181045/15.
DR P-PSDB; AAY05000.
XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
XX
PS Claim 22; Fig 50P; 309pp; French.
XX
CC Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
XX
SQ - Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 other;

Query Match 100.0%; Score 396; DB 20; Length 1068;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGAGTTCGCCATTCGGATC 60
DB 670 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGAGTTCGCCATTCGGATC 729
QY 61 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
DB 730 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 789
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAACGGCGCAGCA 180
DB 790 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAACGGCGCAGCA 849
QY 181 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 850 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 909
QY 241 ATCACCAGGCGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 300
DB 910 ATCACCAGGCGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 969
QY 301 GGGCATATCCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGGCGCGT 360
DB 970 GGGCATATCCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGGCGCGT 1029
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
DB 1030 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1065

RESULT 11
AAX34252
ID AAX34252 standard; DNA; 1143 BP.
XX
AC AAX34252;
XX
DT 06-JUL-1999 (first entry)
XX
DE Mycobacterium species nucleic acid sequence 50F.
XX
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS Mycobacterium sp.
XX
PN W09909186-A2.
XX

```

---

```

PD 25-FEB-1999.
XX
PF 14-AUG-1998; 98WO-FR01813.
XX
PR 11-SEP-1997; 97FR-0011325.
PR 14-AUG-1997; 97FR-0010404.
XX
PA (INSP ) INST PASTEUR.
XX
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
XX
DR WPI: 1999-181045/15.
DR P-PSDB; AAY04830.
XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
XX
PS Claim 22; Fig 50F; 309pp; French.
XX
CC Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
XX
SQ Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 other;

Query Match 100.0%; Score 396; DB 20; Length 1143;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGAGTTCGCCATTCGGATC 60
DB 745 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGAGTTCGCCATTCGGATC 804
QY 61 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
DB 805 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 864
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAACGGCGCAGCA 180
DB 865 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAACGGCGCAGCA 924
QY 181 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 925 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 984
QY 241 ATCACCAGGCGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 300
DB 985 ATCACCAGGCGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 1044
QY 301 GGGCATATCCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGGCGCGT 360
DB 1045 GGGCATATCCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGGCGCGT 1104
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
DB 1105 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1140

RESULT 12
AAL40771
ID AAL40771 standard; DNA; 1742 BP.
XX
AC AAL40771;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ral2-WT1 fusion protein.
XX

```



Db	265	ATCACCGCGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC	324
Qy	301	GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGGACGCGT	360
Db	325	GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGGACGCGT	384
Qy	361	ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC	396
Db	385	ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC	420
RESULT 13			
AA	47084	AAD47084 standard; DNA; 2190 BP.	
XX	XX	AAD47084;	
XX	XX	27-JAN-2003 (first entry)	
DE	XX	Mycobacterium sp. MTB72FmutSA fusion protein encoding DNA.	
XX	XX	Vaccine; immunity; diagnostic agent; gene therapy; TBH9; antigen;	
KW	XX	Ra35MutSA; Ral2; MTB72MutSA; chimeric; gene; ds.	
KW	XX	Chimeric - Mycobacterium sp.	
OS	XX	Chimeric - Mycobacterium tuberculosis.	
XX	XX	Key Location/Qualifiers	
FT	CDS	1..2190	
FT	FT	/*tag= a	
FT	FT	/product= "MTB72FmutSA fusion protein"	
XX	XX	WO200272792-A2.	
PN	XX	19-SEP-2002.	
PD	XX	13-MAR-2002; 2002WO-US08223.	
XX	XX	13-MAR-2001; 2001US-275837P.	
XX	XX	(CORI-) CORIXA CORP.	
PA	XX	Skeiky Y, Brannon M, Guderian J;	
PI	XX	WPI; 2002-759844/82.	
DR	XX	P-PSDB; AAE29709.	
DR	XX	New recombinant nucleic acid molecule comprising a Leishmania TSA,	
PT	XX	Leif, M15 or 6H polynucleotide, useful as vaccine to elicit protective	
PT	XX	immunity against pathogenic microorganisms e.g. Leishmania and	
PT	XX	Mycobacterium tuberculosis	
XX	XX	Disclosure; Page 92-93; 155pp; English.	
PS	XX	The invention relates to a recombinant nucleic acid molecule encoding a	
CC	XX	fusion polypeptide. The recombinant nucleic acid comprises a heterologous	
CC	XX	polynucleotide sequence encoding an antigen or an antigenic fragment from	
CC	XX	Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a	
CC	XX	polypeptide or its fragment. The Leishmania polynucleotide is selected	
CC	XX	from TSA, Leif, M15, and 6H polynucleotides. Sequences of the invention	
CC	XX	are used in methods for eliciting immune response in mammals. They are	
CC	XX	useful as vaccines to elicit protective immunity against pathogenic	
CC	XX	microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion	
CC	XX	polypeptides are used for enhancing the expression of polynucleotides,	
CC	XX	as in vivo diagnostic agents and for raising antibodies in a non-human	
CC	XX	animal. The invention is used in gene therapy. The present sequence is	
CC	XX	a DNA encoding MTB72F fusion protein. This fusion protein comprises	
CC	XX	Ra35MutSA mutant protein and Ral2 protein from M. tuberculosis and	
CC	XX	TBH9 protein from Mycobacterium sp.	
XX	XX	Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 other;	
SQ	XX		

(CORI-) CORIXA CORP.  
Skeiky Y, Reed S, Alderson M;  
WPI: 2002-147798/19.  
P-PSDB: AAE17573.  
Composition comprising MTB39 antigen and MTB32A antigen from  
Mycobacterium species, useful for eliciting immune response in a  
subject -  
Claim 81; Page 108-109; 136pp; English.  
The present invention relates to fusion proteins containing at least  
two Mycobacterium species antigens, nucleotides encoding them and  
compositions comprising such fusion proteins. The present invention  
particularly relates to nucleic acids encoding fusion proteins that  
include two or more individual M. tuberculosis antigens which increase  
the serological sensitivity of sera from individuals infected with  
tuberculosis and methods for their use in diagnosis, prevention and  
treatment of tuberculosis infection. Sequences of the invention are  
useful for eliciting an immune response in a mammal, e.g., human,  
immunised with BCG. They are useful in the diagnosis, treatment and  
prevention of Mycobacterium infection. The fusion proteins and the  
polynucleotides are useful as diagnostic tools in patients infected  
with Mycobacterium, in vitro and in vivo assays for detecting humoral  
antibodies or cell-mediated immunity against M. tuberculosis, for the  
diagnosis of an infection or monitoring of disease progression, as  
immunogens to generate or elicit a protective immune response in a  
patient and for raising anti-M. tuberculosis antibodies in a non-human  
animal. Sequences of the invention are also used as vaccines. MTB32A  
fusion proteins of the invention are useful as in vivo diagnostic agents  
for intradermal skin test. The present sequence is a DNA encoding  
Mycobacterium species MTB72FwutSA (Ral12-TbH9-Ra35WutSA) mutant fusion  
protein.  
Sequence 2190 BP; 340 A; 697 C; 797 G; 356.T; 0 other;

Query Match	100.0%	Score 396;	DB 24;	Length 2190;
Best Local Similarity	100.0%;	Pred. No. 1.3e-84;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 other;				
1	ACGCCGCCGTCCGATACCTTCCAGCTGTCCAGGGTGGGCAGGAGATTCGCCATTCCGATC	60		
22	ACGCCGCCGTCCGATACCTTCCAGCTGTCCAGGGTGGGCAGGAGATTCGCCATTCCGATC	81		

QY	61	GGGAGCGGATGGGGATCGGGGCCAGATCGGATCGGGTGGGGTCAACCCACCGCTTCAT	120
Db	82	GGGAGCGGATGGCGATCGGGGCCAGATCGGATCGGGTGGGGTCAACCCACCGCTTCAT	141
QY	121	ATCGGGCTACCGGCTTCCTCGGCTTGGGTTCGACAAACAGGCAACGGCGCAGCA	180
Db	142	ATCGGGCTACCGGCTTCCTCGGCTTGGGTTCGACAAACAGGCAACGGCGCAGCA	201
QY	181	GTCCAAAGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGGACGTG	240
Db	202	GTCCAAAGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGGACGTG	261
QY	241	ATCACCGCGTGCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAAC	300
Db	262	ATCACCGCGTGCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAAC	321
QY	301	GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCTGGCGGCACGGGT	360

RESULT 15

Db	322	GGGCATCATCCGGTGACGTCATCTCGGTGACCTTGGCAAAACCAAGTCGGCGGCACGCGT	381
Db	361	ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC	396
Db	382	ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC	417

## RESULT 15

AA140773  
ID AAL40773 standard; DNA; 2191 BP.

AC AAL40773;

DT 03-OCT-2002 (first entry)

DE Nucleotide sequence encoding Ral2-H9-32A fusion protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.

XX Chimeric - Unidentified.

XX Key Location/Qualifiers

FT CDS 1..2190

FT /tag- a

FT /product- "Ral2-H9-32A fusion protein"

XX WO200125401-A2;

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27652.

XX 07-OCT-1999; 99US-0158585.

XX (CORI-) CORIXA CORP.

XX Skelky Y, Guderian J;

XX WPI; 2001-266299/27.

XX P-PSDB; AAO22142.

XX Recombinant nucleic acid molecule for producing high yield expression  
PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide

XX Disclosure; Fig 6; 39pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
CC kDa C-terminal fragment of serine protease antigen MTB32A of  
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
CC The recombinant fusion nucleic acids and polypeptides are useful for  
CC providing stable and high yield expression of fusion polypeptides of both  
CC eukaryotic and prokaryotic origin and to encode a protein product for use  
CC as an antigen for detecting serum antibodies. The presence of serum  
CC antibodies to M. tuberculosis antigens in an individual indicates that  
CC the individual is infected with it. The fusion polypeptides are useful as  
CC sources of proteins for monitoring binding of serum antibodies to fusion  
CC proteins and as an immunogen to induce and/or enhance immune responses.  
CC The coding sequences can be ligated with a coding sequence of another  
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
CC represents the DNA encoding the Ral2-H9-32A fusion protein.

XX Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 1 other;

Query Match 100.0%; Score 396; DB 23; Length 2191;

Best Local Similarity 100.0%; Pred. No. 1.3e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGCGGGTGGGCGAGGATTCGCCATTCGCATC 60

DB 22 ACGGCGCGTCCGATTAACCTCCAGCTGCGGGTGGGCGAGGATTCGCCATTCGCATC 81

QY 61 GGGCAGGCGATGCGGATCGGGGCGAGATCCGATCGGGTGGGCGGTCACCCACCGTTCAT 120

DB 82 GGGCAGGCGATGCGGATCGGGGCGAGATCCGATCGGGTGGGCGGTCACCCACCGTTCAT 141

QY 121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTGACAAACACGGCAGCGGACGCA 180  
DB 142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTGACAAACACGGCAGCGGACGCA 201  
QY 181 GTCCAAACGGGTGGTGGGAGCGCTCCGGGCGCAAGTCTCGGCATCTCCACCGGCGAGCTG 240  
DB 202 GTCCAAACGGGTGGTGGGAGCGCTCCGGGCGCAAGTCTCGGCATCTCCACCGGCGAGCTG 261  
QY 241 ATCACCAGCGGTGACGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGCTTAAC 300  
DB 262 ATCACCAGCGGTGACGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGCTTAAC 321  
QY 301 GGGCATCATCCCGGTGACGCTATCTCGGTGACCTGCGAAACCAAGTCGGGCGGACGCGT 360  
DB 322 GGGCATCATCCCGGTGACGCTATCTCGGTGACCTGCGAAACCAAGTCGGGCGGACGCGT 381  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGGACCCCGGCC 396  
DB 382 ACAGGGAACGTGACATTTGGCCGAGGAGGACCCCGGCC 417

Search completed: September 5, 2003, 07:24:26  
Job time : 251 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:04 ; Search time 75 Seconds  
(without alignments)  
2330.503 Million cell updates/sec

Title: US-09-684-215A-3

Perfect score: 396

Sequence: 1 acggccgcgtccgataactt.....tggccgaggagccccggcc 396

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	2287	4	US-09-223-040-1
2	396	100.0	4403765	3	US-09-103-840A-2
3	396	100.0	4411529	3	US-09-103-840A-1
4	394.4	99.6	447	3	US-08-818-112-4
5	394.4	99.6	447	4	US-08-818-111-4
6	394.4	99.6	447	4	US-09-056-556-4
7	394.4	99.6	447	4	US-09-072-596-4
8	391.2	98.8	1872	3	US-08-818-112-17
9	391.2	98.8	1872	4	US-08-818-111-17
10	391.2	98.8	1872	4	US-09-056-556-17
11	391.2	98.8	1872	4	US-09-072-596-17
12	354	89.4	822	4	US-09-736-457-1862
13	354	89.4	900	4	US-09-643-597-353
14	354	89.4	900	4	US-09-606-421B-353
15	354	89.4	945	4	US-09-736-457-1861
16	354	89.4	1012	4	US-09-643-597-351
17	354	89.4	1012	4	US-09-606-421B-351
18	354	89.4	1464	4	US-09-620-412C-348
19	354	89.4	1464	4	US-09-598-419-348
20	354	89.4	1557	4	US-09-620-412C-332
21	354	89.4	1557	4	US-09-598-419-332
22	354	89.4	1578	4	US-09-556-877-188
23	354	89.4	1578	4	US-09-620-412C-188
24	354	89.4	1578	4	US-09-598-419-188
25	354	89.4	1752	4	US-09-620-412C-352
26	354	89.4	1752	4	US-09-598-419-352
27	354	89.4	1752	4	US-09-620-412C-336

28	354	89.4	1758	4	US-09-598-419-336	Sequence 336, App
29	354	89.4	1860	4	US-09-620-412C-308	Sequence 308, App
30	354	89.4	1860	4	US-09-598-419-308	Sequence 308, App
31	354	89.4	1896	4	US-09-620-412C-324	Sequence 324, App
32	354	89.4	1896	4	US-09-598-419-324	Sequence 324, App
33	354	89.4	1941	4	US-09-620-412C-316	Sequence 316, App
34	354	89.4	1941	4	US-09-598-419-316	Sequence 316, App
35	354	89.4	1965	4	US-09-620-412C-340	Sequence 340, App
36	354	89.4	1965	4	US-09-598-419-340	Sequence 340, App
37	354	89.4	2052	4	US-09-620-412C-356	Sequence 356, App
38	354	89.4	2052	4	US-09-598-419-356	Sequence 356, App
39	354	89.4	2076	4	US-09-620-412C-312	Sequence 312, App
40	354	89.4	2076	4	US-09-598-419-312	Sequence 312, App
41	354	89.4	2103	4	US-09-620-412C-344	Sequence 344, App
42	354	89.4	2103	4	US-09-598-419-344	Sequence 344, App
43	354	89.4	2148	4	US-09-620-412C-320	Sequence 320, App
44	354	89.4	2148	4	US-09-598-419-320	Sequence 320, App
45	354	89.4	2148	4	US-09-620-412C-328	Sequence 328, App
					US-09-598-419-320	Sequence 320, App

#### ALIGNMENTS

#### RESULT 1

US-09-223-040-1

; Sequence 1, Application US/092223040

; Patent No. 6544522

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir

; APPLICANT: Alderson, Mark

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; FILE REFERENCE: 014058-0090100S

; CURRENT APPLICATION NUMBER: US/09/223,040

; CURRENT FILING DATE: 1998-12-30

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2287

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion

; OTHER INFORMATION: protein Ra12-TbH9-Ra35

; NAME/KEY: modified\_base

; LOCATION: (30)

; OTHER INFORMATION: n = g, a, c o r t

; NAME/KEY: modified\_base

; LOCATION: (33)

; OTHER INFORMATION: n = g, a, c o r t

; NAME/KEY: CDS

; LOCATION: (42)..(2231)

; NAME/KEY: modified\_base

; LOCATION: (2270)

; OTHER INFORMATION: n = g, a, c o r t

US-09-223-040-1

Query Match 100.0%; Score 396; DB 4; Length 2287;

Best Local Similarity 100.0%; Pred. No. 2e-91;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCCCGATC 60

|||||

Db 63 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCCCGATC 122

|||||

Qy 61 GGGCAGCGATGCGGATCGCGGGCGAGATCGGATCGGTTGGGGGTCAACCACCGTTTCAT 120

|||||

Db 123 GGGCAGCGATGCGGATCGCGGGCGAGATCGGATCGGTTGGGGGTCAACCACCGTTTCAT 182

|||||

Qy 121 ATCCGGGCTACCGCTTCCCTCGGTTGGGTTTGTGCACACACACGCGGCGCAGCA 180

|||||

```
Db 183 ATCGGGCTACCGCTTCCTCGGCTGGGTGTTGTGCACAACACGGCAACGGCGCACGA 242
QY 181 GTCCAAACGGTGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGGTG 240
Db 243 GTCCAAACGGTGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGGTG 302
QY 241 ATCACCGCGTGCAGCGGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 300
Db 303 ATCACCGCGTGCAGCGGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 362
QY 301 GGCATCATCCCGGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACGGCT 360
Db 363 GGCATCATCCCGGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACGGCT 422
QY 361 ACAGGAACGTGATTCGGCGGAGGACCCCGGCC 396
Db 423 ACAGGAACGTGATTCGGCGGAGGACCCCGGCC 458

RESULT 2
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 100.0%; Score 396; DB 3; Length 4403765;
Best Local Similarity 100.0%; Pred. No. 7e-91;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 1 ACGGGCGGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60
Db 151984 ACGGGCGGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 152043
QY 61 GGGCAGCGGATGGCGATCGCGGCCAGATCCCGATCGGGTGGGGGTTCACCCACCGTTTAT 120
Db 152044 GGGCAGCGGATGGCGATCGCGGCCAGATCCCGATCGGGTGGGGGTTCACCCACCGTTTAT 152103
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTTGTGCACAACACGGCAACGGCGACGA 180
Db 152104 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTTGTGCACAACACGGCAACGGCGACGA 152163
QY 181 GTCCAAACGGTGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 240
Db 152164 GTCCAAACGGTGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 152223
QY 241 ATCACCGCGTGCAGCGGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 300
Db 152224 ATCACCGCGTGCAGCGGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 152283
QY 301 GGCATCATCCCGGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACGGCT 360
Db 152284 GGCATCATCCCGGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACGGCT 152343
```

```
QY 361 ACAGGAACGTGATTCGGCGGAGGACCCCGGCC 396
Db 152344 ACAGGAACGTGATTCGGCGGAGGACCCCGGCC 152379

RESULT 3
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 100.0%; Score 396; DB 3; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 7e-91;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGGCGGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60
Db 151815 ACGGGCGGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 151874
QY 61 GGGCAGCGGATGGCGATCGCGGCCAGATCCCGATCGGGTGGGGGTTCACCCACCGTTTAT 120
Db 151875 GGGCAGCGGATGGCGATCGCGGCCAGATCCCGATCGGGTGGGGGTTCACCCACCGTTTAT 151934
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTTGTGCACAACACGGCAACGGCGACGA 180
Db 151935 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTTGTGCACAACACGGCAACGGCGACGA 151994
QY 181 GTCCAAACGGTGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 240
Db 151995 GTCCAAACGGTGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 152054
QY 241 ATCACCGCGGTCGAGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 300
Db 152055 ATCACCGCGGTCGAGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 152114
QY 301 GGCATCATCCCGGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACGGCT 360
Db 152115 GGCATCATCCCGGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACGGCT 152174
QY 361 ACAGGAACGTGATTCGGCGGAGGACCCCGGCC 396
Db 152175 ACAGGAACGTGATTCGGCGGAGGACCCCGGCC 152210

RESULT 4
US-08-818-112-4
; Sequence 4, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
```

;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
;; NUMBER OF INVENTIONS: AND DIAGNOSIS OF TUBERCULOSIS  
;; NUMBER OF SEQUENCES: 153  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SEED AND BERRY LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98104-7092  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/818,112  
;; FILING DATE: 13-MAR-1997  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maki, David J.  
;; REGISTRATION NUMBER: 210121.411C6  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-818-112-4

Query Match 99.6%; Score 394.4; DB 3; Length 447;  
Best Local Similarity 99.7%; Pred. No. 3.8e-91;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTGCGCATTCGATC 60  
DB |||||||  
QY 11 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTGCGCATTCGATC 70  
DB |||||||  
QY 61 GGCAGCGGATGCGATCGCGGCGAGATCCGATCGGGTGGGAGGATTGCGCATTCGATC 120  
DB |||||||  
QY 71 GGCAGCGGATGCGATCGCGGCGAGATCCGATCGGGTGGGAGGATTGCGCATTCGATC 130  
DB |||||||  
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 180  
DB |||||||  
QY 131 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 190  
DB |||||||  
QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGAGCTG 240  
DB |||||||  
QY 191 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGAGCTG 250  
DB |||||||  
QY 241 ATCCACCGGTCGACGCGCTCGATCAACTCGGCGACCGCGATGCGGCGGCTTAAC 300  
DB |||||||  
QY 251 ATCCACCGGTCGACGCGCTCGATCAACTCGGCGACCGCGATGCGGCGGCTTAAC 310  
DB |||||||  
QY 301 GGCATCATCCCGGTGACGTCATCTCGGTGACCTGCGCAACGCGGCGGCGGCGGT 360  
DB |||||||  
QY 311 GGCATCATCCCGGTGACGTCATCTCGGTGACCTGCGCAACGCGGCGGCGGCGGT 370  
DB |||||||  
QY 361 ACAGGAAAGTGACATTGGCCGAGGACCCCGGCC 396  
DB |||||||  
QY 371 ACAGGAAAGTGACATTGGCCGAGGACCCCGGCC 406  
DB |||||||

## RESULT 5

US-08-818-111-4  
; Sequence 4, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.

;; APPLICANT: Skeiky, Yasir A.W.  
;; APPLICANT: Dillon, Davin C.  
;; APPLICANT: Campos-Neto, Antonia  
;; APPLICANT: Houghton, Raymond  
;; APPLICANT: Teardzik, Thomas S.  
;; APPLICANT: Twardzik, Daniel R.  
;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
;; NUMBER OF SEQUENCES: 148  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SEED AND BERRY LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98104-7092  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/818,111  
;; FILING DATE: 13-MAR-1997  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maki, David J.  
;; REGISTRATION NUMBER: 210121.417C6  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-818-111-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
Best Local Similarity 99.7%; Pred. No. 3.8e-91;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTGCGCATTCGATC 60  
DB |||||||  
QY 11 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTGCGCATTCGATC 70  
DB |||||||  
QY 61 GGCAGCGGATGCGATCGCGGCGAGATCCGATCGGGTGGGAGGATTGCGCATTCGATC 120  
DB |||||||  
QY 71 GGCAGCGGATGCGATCGCGGCGAGATCCGATCGGGTGGGAGGATTGCGCATTCGATC 130  
DB |||||||  
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 180  
DB |||||||  
QY 131 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACGCGGCGACGCA 190  
DB |||||||  
QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGAGCTG 240  
DB |||||||  
QY 191 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGAGCTG 250  
DB |||||||  
QY 241 ATCCACCGGTCGACGCGCTCGATCAACTCGGCGACCGCGATGCGGCGGCTTAAC 300  
DB |||||||  
QY 251 ATCCACCGGTCGACGCGCTCGATCAACTCGGCGACCGCGATGCGGCGGCTTAAC 310  
DB |||||||  
QY 301 GGCATCATCCCGGTGACGTCATCTCGGTGACCTGCGCAACGCGGCGGCGGCGGT 360  
DB |||||||  
QY 311 GGCATCATCCCGGTGACGTCATCTCGGTGACCTGCGCAACGCGGCGGCGGCGGT 370  
DB |||||||  
QY 361 ACAGGAAAGTGACATTGGCCGAGGACCCCGGCC 396  
DB |||||||  
QY 371 ACAGGAAAGTGACATTGGCCGAGGACCCCGGCC 406  
DB |||||||

## RESULT 6

US-09-056-556-4  
; Sequence 4, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.457  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-056-556-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
Best Local Similarity 99.7%; Pred. No. 3.8e-91;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ACGGCCGGTCCGATAACTTCCAGCTGTC	CGAGGTGGCGAGGATTCCGCATTC	CGGATC	60
DB	11	ACGGCCGGTCCGATAACTTCCAGCTGTC	CGAGGTGGCGAGGATTCCGCATTC	CGGATC	70
QY	61	GGCAGGCGATGGCGATCGCGGCGCAGAT	CGCGGCGCAGATCGCGGCGCAGAT	CGCGGCGCAGAT	120
DB	71	GGCAGGCGATGGCGATCGCGGCGCAGAT	CGCGGCGCAGATCGCGGCGCAGAT	CGCGGCGCAGAT	130
QY	121	ATCGGGCTACCGCTTCTCGGCTTGGGT	GTTCGACAAACAGCGGCGCAGCA	180	
DB	131	ATCGGGCTACCGCTTCTCGGCTTGGGT	GTTCGACAAACAGCGGCGCAGCA	190	
QY	181	GTCCAAACCGGTGGTGGGAGCGCTCCCG	CGCGGCGCAGTCTCGGCGCAGTCT	CGGCGGCGCAGTCT	240
DB	191	GTCCAAACCGGTGGTGGGAGCGCTCCCG	CGCGGCGCAGTCTCGGCGCAGTCT	CGGCGGCGCAGTCT	250
QY	181	GTCCAAACCGGTGGTGGGAGCGCTCCCG	CGCGGCGCAGTCTCGGCGCAGTCT	CGGCGGCGCAGTCT	240
DB	191	GTCCAAACCGGTGGTGGGAGCGCTCCCG	CGCGGCGCAGTCTCGGCGCAGTCT	CGGCGGCGCAGTCT	250
QY	241	ATCACCAGCGGTGGCGGCGCTCCGATCA	ACTTCCGCGGCGCAGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	300
DB	251	ATCACCAGCGGTGGCGGCGCTCCGATCA	ACTTCCGCGGCGCAGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	310
QY	301	GGGATCATCCCGGTGAGTGGCGGAGG	ACCCCGGCGGCGGCGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	360
DB	371	ACAGGGAACGTGACATTGGCGGAGG	ACCCCGGCGGCGGCGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	406

RESULT 7  
US-09-072-596-4  
; Sequence 4, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-072-596-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
Best Local Similarity 99.7%; Pred. No. 3.8e-91;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ACGGCCGGTCCGATAACTTCCAGCTGTC	CGAGGTGGCGAGGATTCCGCATTC	CGGATC	60
DB	11	ACGGCCGGTCCGATAACTTCCAGCTGTC	CGAGGTGGCGAGGATTCCGCATTC	CGGATC	70
QY	61	GGCAGGCGATGGCGATCGCGGCGCAGAT	CGCGGCGCAGATCGCGGCGCAGAT	CGCGGCGCAGAT	120
DB	71	GGCAGGCGATGGCGATCGCGGCGCAGAT	CGCGGCGCAGATCGCGGCGCAGAT	CGCGGCGCAGAT	130
QY	121	ATCGGGCTACCGCTTCTCGGCTTGGGT	GTTCGACAAACAGCGGCGCAGCA	180	
DB	131	ATCGGGCTACCGCTTCTCGGCTTGGGT	GTTCGACAAACAGCGGCGCAGCA	190	
QY	181	GTCCAAACCGGTGGTGGGAGCGCTCCCG	CGCGGCGCAGTCTCGGCGCAGTCT	CGGCGGCGCAGTCT	240
DB	191	GTCCAAACCGGTGGTGGGAGCGCTCCCG	CGCGGCGCAGTCTCGGCGCAGTCT	CGGCGGCGCAGTCT	250
QY	241	ATCACCAGCGGTGGCGGCGCTCCGATCA	ACTTCCGCGGCGCAGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	300
DB	251	ATCACCAGCGGTGGCGGCGCTCCGATCA	ACTTCCGCGGCGCAGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	310
QY	301	GGGATCATCCCGGTGAGTGGCGGAGG	ACCCCGGCGGCGGCGGCGGCGGCGG	CGGCGGCGGCGGCGGCGG	360



Db 311 GGGCATCATCCCGTGTACGTATCTCGGTGAATCGGCAACAAAGTCGGGGCGGACGCGT 370  
Qy 361 ACAGGGAACGTGACATTGGCGAGGAGCCCGCGGCC 396  
Db 371 ACAGGGAACGTGACATTGGCGAGGAGCCCGCGGCC 406

RESULT 8  
US-08-818-112-17  
; Sequence 17, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-112-17

Query Match 98.8%; Score 391.2; DB 3; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 758 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 817  
Qy 61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 120  
Db 818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 877  
Qy 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTTCGACACACACGCGGACGACGTA 180  
Db 878 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTTCGACACACACGCGGACGACGTA 937  
Qy 181 GTCCAAACGCGTGGTGGGAGGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGAGTG 240  
Db 938 GTCCAAACGCGTGGTGGGAGGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGAGTG 997

Qy 241 ATCACCGCGTGTACGGCGCTCCGATCAATCGGCCACCGCGATGGGGAGCGCGCTTAAC 300  
Db 998 ATCACCGCGTGTACGGCGCTCCGATCAATCGGCCACCGCGATGGGGAGCGCGCTTAAC 1057  
Qy 301 GGGCATCATCCCGTGTACGTATCTCGGTGAATCGGCAACAAAGTCGGGGCGGACGCGT 360  
Db 1058 GGGCATCATCCCGTGTACGTATCTCGGTGAATCGGCAACAAAGTCGGGGCGGACGCGT 1117  
Qy 361 ACAGGGAACGTGACATTGGCGAGGAGCCCGCGGCC 396  
Db 1118 ACAGGGAACGTGACATTGGCGAGGAGCCCGCGGCC 1153

RESULT 9  
US-08-818-111-17  
; Sequence 17, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,111  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-111-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 758 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 817  
Qy 61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 120  
Db 818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 877  
Qy 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTTCGACACACACGCGGACGACGTA 180  
Db 878 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTTCGACACACACGCGGACGACGTA 937

QY 181 GTCCAAACGGGTGGTCCGGAGCGCTCCGGGGCAAGTCTCGGATCTCCACCGCGGACGTG 240  
|||||  
Db 938 GTCCAAACGGGTGGTCCGGAGCGCTCCGGGGCAAGTCTCGGATCTCCACCGCGGACGTG 997  
|||||  
QY 241 ATCACCAGCGGTGGTCCGGAGCGCTCCGGATCAACTCGGCGGCGGATCGGCGGCGGCTTAAC 300  
Db 998 ATCACCAGCGGTGGTCCGGAGCGCTCCGGATCAACTCGGCGGCGGATCGGCGGCGGCTTAAC 1057  
QY 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCCGGGGCGGCGGCT 360  
Db 1058 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCCGGGGCGGCGGCT 1117  
QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGGCC 396  
Db 1118 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGGCC 1153

RESULT 10  
US-09-056-556-17  
; Sequence 17, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.457  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-056-556-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACAGCGCGGTCCGATAACTTCCAGCTGTCCAGGCGGCGGAGGATTCGCCATTCGGATC 60  
|||||  
Db 758 ACAGCGCGGTCCGATAACTTCCAGCTGTCCAGGCGGCGGAGGATTCGCCATTCGGATC 817  
QY 61 GGGCAGGCGATGGCGATCGCGGCGGCGGATCGATCGGTGGGGGTCAACCCACCGCTTAT 120  
|||||  
Db 818 GGGCAGGCGATGGCGATCGCGGCGGCGGATCGATCGGTGGGGGTCAACCCACCGCTTAT 877  
QY 121 ATCGGGCGCTACGGCTTCTCGGCTTGGTGTGTCGACACCAACGCAACGGCGGCGACGA 180  
|||||

Db 878 ATCGGGCGCTACGGCTTCTCGGCTTGGTGTGTCGACACCAACGCAACGGCGGCGACGA 937  
QY 181 GTCCAAACGGGTGGTCCGGAGCGCTCCGGGGCAAGTCTCGGATCTCCACCGCGGACGTG 240  
|||||  
Db 938 GTCCAAACGGGTGGTCCGGAGCGCTCCGGGGCAAGTCTCGGATCTCCACCGCGGACGTG 997  
|||||  
QY 241 ATCACCAGCGGTGGTCCGGAGCGCTCCGGATCAACTCGGCGGCGGATCGGCGGCGGCTTAAC 300  
Db 998 ATCACCAGCGGTGGTCCGGAGCGCTCCGGATCAACTCGGCGGCGGATCGGCGGCGGCTTAAC 1057  
QY 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCCGGGGCGGCGGCT 360  
Db 1058 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCCGGGGCGGCGGCT 1117  
QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGGCC 396  
Db 1118 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGGCC 1153

RESULT 11  
US-09-072-596-17  
; Sequence 17, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-072-596-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACAGCGCGGTCCGATAACTTCCAGCTGTCCAGGCGGCGGAGGATTCGCCATTCGGATC 60  
|||||  
Db 758 ACAGCGCGGTCCGATAACTTCCAGCTGTCCAGGCGGCGGAGGATTCGCCATTCGGATC 817

Oy 61 GGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 818 GGCAGCGATGGCGATCGCGGGCCAAATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 877  
Oy 121 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTGTCGACAAACAGCGGCGGACGA 180  
Db 878 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTGTCGACAAACAGCGGCGGACGA 937  
Oy 181 GTCCAAACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGACGTG 240  
Db 938 GTCCAAACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGACGTG 997  
Oy 241 ATCCAGCGGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCTTAAAC 300  
Db 998 ATCCAGCGGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCTTAAAC 1057  
Oy 301 GGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGACCGGT 360  
Db 1058 GGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGACCGGT 1117  
Oy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

## RESULT 12

US-09-736-457-1862  
; Sequence 1862, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736.457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 1862  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1862

Query Match 89.4%; Score 354; DB 4; Length 822;  
Best Local Similarity 95.7%; Pred. No. 6.8e-81;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
Oy 1 ACGCCGCGTCCGATTAACCTCCAGCTGCCAGGTTGGGAGGATTCGCCATTCCGATC 60  
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGCCAGGTTGGGAGGATTCGCCATTCCGATC 81  
Oy 61 GGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
Oy 121 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTGTCGACAAACAGCGGCGGACGA 180  
Db 130 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTGTCGACAAACAGCGGCGGACGA 189  
Oy 181 GTCCAAACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGACGTG 240  
Db 190 GTCCAAACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGACGTG 249

Oy 241 ATCCAGCGGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCCAGCGGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 309  
Oy 301 GGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGACCGGT 360  
Db 310 GGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGACCGGT 369  
Oy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 405

## RESULT 13

US-09-643-597-353  
; Sequence 353, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643.597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-643-597-353

Query Match 89.4%; Score 354; DB 4; Length 900;  
Best Local Similarity 95.7%; Pred. No. 6.9e-81;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
Oy 1 ACGCCGCGTCCGATTAACCTCCAGCTGCCAGGTTGGGAGGATTCGCCATTCCGATC 60  
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGCCAGGTTGGGAGGATTCGCCATTCCGATC 81  
Oy 61 GGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
Oy 121 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTGTCGACAAACAGCGGCGGACGA 180  
Db 130 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTGTCGACAAACAGCGGCGGACGA 189  
Oy 181 GTCCAAACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGACGTG 240  
Db 190 GTCCAAACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGACGTG 249  
Oy 241 ATCCAGCGGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCCAGCGGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 309  
Oy 301 GGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGACCGGT 360  
Db 310 GGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGACCGGT 369  
Oy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 405

Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405

## RESULT 14

US-09-606-421B-353  
; Sequence 353, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aljun

; APPLICANT: Skeiky, Yasir A.W

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606,421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 353

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-606-421B-353

Query Match 89.4%; Score 354; DB 4; Length 900;

Best Local Similarity 95.7%; Pred. No. 6.9e-81;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGAGGTGGCAGGGATTCCGCCATTCGGATC 60

Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGAGGTGGCAGGGATTCCGCCATTCGGATC 81

QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 120

Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATC-----AAGCTTCCCACCGTTTCA 129

QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGGCAGCA 180

Db 130 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGGCAGCA 189

QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCAGCTG 240

Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCAGCTG 249

QY 241 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGTTAAC 300

Db 250 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGTTAAC 309

QY 301 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGCAGCGGT 360

Db 310 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGCAGCGGT 369

QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 396

Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405

## RESULT 15

US-09-736-457-1861

; Sequence 1861, Application US/09736457

; Patent No. 6509448

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aljun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457

; CURRENT FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 1864

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1861

; LENGTH: 945

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-736-457-1861

Query Match 89.4%; Score 354; DB 4; Length 945;

Best Local Similarity 95.7%; Pred. No. 7e-81;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGAGGTGGCAGGGATTCCGCCATTCGGATC 60

Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGAGGTGGCAGGGATTCCGCCATTCGGATC 81

QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 120

Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATC-----AAGCTTCCCACCGTTTCA 129

QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGGCAGCA 180

Db 130 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGGCAGCA 189

QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCAGCTG 240

Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCAGCTG 249

QY 241 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGTTAAC 300

Db 250 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGTTAAC 309

QY 301 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGCAGCGGT 360

Db 310 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGCAGCGGT 369

QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 396

Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405

Search completed: September 5, 2003, 08:28:53

Job time : 87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:04 ; Search time 1535 Seconds  
(without alignments)  
593.349 Million cell updates/sec

Title: US-09-684-215A-3

Perfect score: 396

Sequence: 1 acggcgcgtccgataactt.....tgccgagggaccgccgccc 396

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1537136 seqs, 114998732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	702	9	US-09-287-849-27
2	396	100.0	702	12	US-10-359-460-27
3	396	100.0	1068	10	US-09-712-363-15
4	396	100.0	2287	9	US-09-287-849-1
5	396	100.0	2287	12	US-10-359-460-1
6	394.4	99.6	447	12	US-10-084-843-4
7	394.4	99.6	447	12	US-10-193-002-4
8	391.2	98.8	1872	12	US-10-084-843-17
9	391.2	98.8	1872	12	US-10-193-002-17
10	354	89.4	675	9	US-09-759-143-822
11	354	89.4	675	9	US-09-780-669-822
12	354	89.4	675	9	US-09-822-827-822
13	354	89.4	675	10	US-09-895-793-822
14	354	89.4	675	10	US-09-895-814-822
15	354	89.4	675	12	US-10-144-678A-822
16	354	89.4	675	13	US-10-012-896-822

17	354	89.4	822	10	US-09-736-457-1862	Sequence 1862, Ap
18	354	89.4	822	10	US-09-902-941-1862	Sequence 1862, Ap
19	354	89.4	822	10	US-09-849-626-1862	Sequence 1862, Ap
20	354	89.4	822	14	US-10-017-754-1862	Sequence 1862, Ap
21	354	89.4	900	9	US-09-735-705-353	Sequence 353, App
22	354	89.4	900	10	US-09-850-716A-353	Sequence 353, App
23	354	89.4	900	10	US-09-897-778-353	Sequence 353, App
24	354	89.4	900	12	US-10-117-982-353	Sequence 353, App
25	354	89.4	915	9	US-09-759-143-834	Sequence 834, App
26	354	89.4	915	9	US-09-780-669-834	Sequence 834, App
27	354	89.4	915	9	US-09-822-827-834	Sequence 834, App
28	354	89.4	915	10	US-09-895-793-834	Sequence 834, App
29	354	89.4	915	10	US-09-895-814-834	Sequence 834, App
30	354	89.4	915	12	US-10-144-678A-834	Sequence 834, App
31	354	89.4	915	13	US-10-012-896-834	Sequence 834, App
32	354	89.4	945	10	US-09-736-457-1861	Sequence 1861, Ap
33	354	89.4	945	10	US-09-902-941-1861	Sequence 1861, Ap
34	354	89.4	945	10	US-09-849-626-1861	Sequence 1861, Ap
35	354	89.4	945	14	US-10-017-754-1861	Sequence 1861, Ap
36	354	89.4	1012	9	US-09-735-705-351	Sequence 351, App
37	354	89.4	1012	10	US-09-850-716A-351	Sequence 351, App
38	354	89.4	1012	10	US-09-897-778-351	Sequence 351, App
39	354	89.4	1012	12	US-10-117-982-351	Sequence 351, App
40	354	89.4	1035	9	US-09-922-217-1084	Sequence 1084, Ap
41	354	89.4	1035	10	US-09-833-263-1084	Sequence 1084, Ap
42	354	89.4	1035	11	US-09-938-864-388	Sequence 388, App
43	354	89.4	1035	13	US-10-025-380-1084	Sequence 1084, Ap
44	354	89.4	1035	14	US-10-125-635A-388	Sequence 388, App
45	354	89.4	1035	14	US-10-002-603-388	Sequence 388, App

## ALIGNMENTS

RESULT 1  
US-09-287-849-27  
; Sequence 27, Application US/09287849  
; Patent No. US20020009459A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 702  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion  
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading  
; OTHER INFORMATION: frame 1  
; NAME/KEY: CDS  
; LOCATION: (1)..(693)  
; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),

OTHER INFORMATION: reading frame 1  
NAME/KEY: CDS  
LOCATION: (2)..(700)  
OTHER INFORMATION: reading frame 2  
NAME/KEY: CDS  
LOCATION: (3)..(701)  
OTHER INFORMATION: reading frame 3  
US-09-287-849-27

Query Match 100.0%; Score 396; DB 9; Length 702;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104; Indels 0; Gaps 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 60  
DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 84  
QY 61 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120  
DB 85 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 144  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGCACGA 180  
DB 145 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGCACGA 204  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 240  
DB 205 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 264  
QY 241 ATCCAGGCGTGGCGGCGCTCCGATCACTCGGGCCACCGGATCGGGCGCGCTTAAAC 300  
DB 265 ATCCAGGCGTGGCGGCGCTCCGATCACTCGGGCCACCGGATCGGGCGCGCTTAAAC 324  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGAGCTGGCAACCAAGTCTCGGCGGCGACGTG 360  
DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGAGCTGGCAACCAAGTCTCGGCGGCGACGTG 384  
QY 361 ACAGGAGCGTGACATTTGGCGGAGGACCCCGGCC 396  
DB 385 ACAGGAGCGTGACATTTGGCGGAGGACCCCGGCC 420

## RESULT 2

US-10-359-460-27  
Sequence 27, Application US/10359460  
Publication No. US2003014791A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skelky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Alderson, Mark  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/10/359,460  
CURRENT FILING DATE: 2003-02-05  
PRIOR APPLICATION NUMBER: US/09/287,849  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 27

LENGTH: 702  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: bi-fusion  
OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading  
OTHER INFORMATION: frame 1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(693)  
OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),  
OTHER INFORMATION: reading frame 1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)..(700)  
OTHER INFORMATION: reading frame 2  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)..(701)  
OTHER INFORMATION: reading frame 3  
US-10-359-460-27

Query Match 100.0%; Score 396; DB 12; Length 702;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104; Indels 0; Gaps 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 60  
DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 84  
QY 61 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120  
DB 85 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 144  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGCACGA 180  
DB 145 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGCACGA 204  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 240  
DB 205 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 264  
QY 241 ATCCAGGCGTGGCGGCGCTCCGATCACTCGGGCCACCGGATCGGGCGCGCTTAAAC 300  
DB 265 ATCCAGGCGTGGCGGCGCTCCGATCACTCGGGCCACCGGATCGGGCGCGCTTAAAC 324  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGAGCTGGCAACCAAGTCTCGGCGGCGACGTG 360  
DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGAGCTGGCAACCAAGTCTCGGCGGCGACGTG 384  
QY 361 ACAGGAGCGTGACATTTGGCGGAGGACCCCGGCC 396  
DB 385 ACAGGAGCGTGACATTTGGCGGAGGACCCCGGCC 420

## RESULT 3

US-09-712-363-15  
Sequence 15, Application US/09712363  
Patent No. US20020164588A1  
GENERAL INFORMATION:  
APPLICANT: Elsenberg, David  
APPLICANT: Rotstein, Sergio H.  
APPLICANT: Marcotte, Edward M.  
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
FILE REFERENCE: 07419-032001  
CURRENT APPLICATION NUMBER: US/09/712,363  
CURRENT FILING DATE: 2000-11-13  
PRIOR APPLICATION NUMBER: PCT/US00/02246  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,531  
PRIOR FILING DATE: 2000-02-01  
PRIOR APPLICATION NUMBER: 60/117,844

```
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-15

Query Match      100.0%; Score 396; DB 10; Length 1068;
Best Local Similarity 100.0%; Pred. No. 7.1e-104;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCCGCTCCGATACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCCGATC 60
    |||||||
Db 670 ACGCCCGCTCCGATACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCCGATC 729

Qy 61 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCGCTTCAT 120
    |||||||
Db 730 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCGCTTCAT 789

Qy 121 ATCGGGCTACCGCTCTCCGCTTGGGTGTCGACAAACAAACGCGGCGACGA 180
    |||||||
Db 790 ATCGGGCTACCGCTCTCCGCTTGGGTGTCGACAAACAAACGCGGCGACGA 849

Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 240
    |||||||
Db 850 GTCCAAACCGGTGTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 909

Qy 241 ATACCCCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGGATGGGCGGCGCTTAAC 300
    |||||||
Db 910 ATACCCCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGGATGGGCGGCGCTTAAC 969

Qy 301 GGGCATATCCCGTGACGTCTCGGTGACCTGGCAACCAAGTGGGCGGCGACCGGT 360
    |||||||
Db 970 GGGCATATCCCGTGACGTCTCGGTGACCTGGCAACCAAGTGGGCGGCGACCGGT 1029

Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
    |||||||
Db 1030 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1065

RESULT 4
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. US20020094591
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
```

```
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; OTHER INFORMATION: protein Rai2-TbH9-Ra35 (designated Mtb32-Mtb39)
; OTHER INFORMATION: fusion)
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

Query Match      100.0%; Score 396; DB 9; Length 2287;
Best Local Similarity 100.0%; Pred. No. 7.1e-104;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCCGCTCCGATACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCCGATC 60
    |||||||
Db 63 ACGCCCGCTCCGATACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCCGATC 122

Qy 61 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCGCTTCAT 120
    |||||||
Db 123 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCGCTTCAT 182

Qy 121 ATCGGGCTACCGCTTCCGCTTGGGTGTCGACAAACAAACGCGGCGACGA 180
    |||||||
Db 183 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTCGACAAACAAACGCGGCGACGA 242

Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 240
    |||||||
Db 243 GTCCAAACCGGTGTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 302

Qy 241 ATACCCCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGGATGGGCGGCGCTTAAC 300
    |||||||
Db 303 ATACCCCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGGATGGGCGGCGCTTAAC 362

Qy 301 GGGCATATCCCGTGACGTCTCGGTGACCTGGCAACCAAGTGGGCGGCGACCGGT 360
    |||||||
Db 363 GGGCATATCCCGTGACGTCTCGGTGACCTGGCAACCAAGTGGGCGGCGACCGGT 422

Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
    |||||||
Db 423 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 458

RESULT 5
US-10-359-460-1
; Sequence 1, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
```

APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/10/359,460  
CURRENT FILING DATE: 2003-02-05  
PRIOR APPLICATION NUMBER: US/09/287,849  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2287  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39)  
OTHER INFORMATION: fusion)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (30)  
OTHER INFORMATION: n = g, a, c o r t  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (33)  
OTHER INFORMATION: n = g, a, c o r t  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (42)..(2231)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (2270)  
OTHER INFORMATION: n = g, a, c o r t  
US-10-359-460-1  
Query Match 100.0%; Score 396; DB 12; Length 2287;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGGGCGGTCGGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 60  
DB 63 ACGGGCGGTCGGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 122  
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCAGTGGGTGGGGTCCACCCACCTTCAAT 120  
DB 123 GGGCAGGCGATGGCGATCGCGGCCAGATCCAGTGGGTGGGGTCCACCCACCTTCAAT 182  
QY 121 ATCGGGCTACCGCCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGACGCGACGCA 180  
DB 183 ATCGGGCTACCGCCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGACGCGACGCA 242  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGACGCGT 240  
DB 243 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGACGCGT 302  
QY 241 ATCAGCGGGTCCAGCGGCTCCGATCACTCGGCCACCGCGATCGCGGACCGGCTTAAC 300  
DB 303 ATCAGCGGGTCCAGCGGCTCCGATCACTCGGCCACCGCGATCGCGGACCGGCTTAAC 362  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACCAAGTCCGGCGGCGACGCGT 360  
DB 363 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACCAAGTCCGGCGGCGACGCGT 422

QY 361 ACAGGAACGACGACATGGCCGAGGAGGACCCCGGCC 396  
DB 423 ACAGGAACGACGACATGGCCGAGGAGGACCCCGGCC 458  
RESULT 6  
US-10-084-843-4  
Sequence 4, Application US/10084843  
Publication No. US20030143243A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
Skelley, Yasir A.W.  
Dillon, Davin C.  
Campos-Neto, Antonio  
Houghton, Raymond  
Vedvick, Thomas S.  
Twardzik, Daniel R.  
Lodes, Michael J.  
Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/084,843  
FILING DATE: 25-FEB-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447; base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-084-843-4  
Query Match 99.6%; Score 394.4; DB 12; Length 447;  
Best Local Similarity 99.7%; Pred. No. 2e-103;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACGGGCGGTCGGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 60  
DB 11 ACGGGCGGTCGGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCAGTGGGTGGGGTCCACCCACCTTCAAT 120  
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCAGTGGGTGGGGTCCACCCACCTTCAAT 130  
QY 121 ATCGGGCTACCGCCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGACGCGACGCA 180  
DB 131 ATCGGGCTACCGCCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGACGCGACGCA 190



QY 181 GTCCAAACGCGTGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGCGT 240  
Db 191 GTCCAAACGCGTGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGCGT 250  
QY 241 ATACACCGGTCGACGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300  
Db 251 ATACACCGGTCGACGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 310  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 360  
Db 311 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 370  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396  
Db 371 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 406

## RESULT 7

US-10-193-002-4  
; Sequence 4, Application US/10193002  
; Publication No. US20030135026A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; Skeiky, Yasir A.W.  
; Dillon, David C.  
; Campos-Neto, Antonio  
; Houghton, Raymond  
; Vedvick, Thomas S.  
; Twardzik, Daniel R.  
; Lodes, Michael J.  
; Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; TUBERCULOSIS  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/193,002  
; FILING DATE: 10-Jul-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-193-002-4

Query Match 99.6%; Score 394.4; DB 12; Length 447;  
Best Local Similarity 99.7%; Pred. No. 2e-103;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGCCCGCGTCCGATACACTTCCAGCTGTCCAGGGTGGGAGGATTCGCATTCCGATC 60  
Db 11 ACGCGCGCGTCCGATACACTTCCAGCTGTCCAGGGTGGGAGGATTCGCATTCCGATC 70  
QY 61 GGGCAGCGGATGGGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCGGCTTCAAT 120  
Db 71 GGGCAGCGGATGGGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCGGCTTCAAT 130  
QY 121 ATCGGGGCTACCGCTTCCCTCGCTGGGTGTTGTCGACAAACGCAACGGCGGCGACGA 180  
Db 131 ATCGGGGCTACCGCTTCCCTCGCTGGGTGTTGTCGACAAACGCAACGGCGGCGACGA 190  
QY 181 GTCCAAACGCGTGGTTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 191 GTCCAAACGCGTGGTTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
QY 241 ATCACCGCGTGGTTCGGGAGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300  
Db 251 ATCACCGCGTGGTTCGGGAGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 310  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 360  
Db 311 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 370  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396  
Db 371 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 406

## RESULT 8

US-10-084-843-17  
; Sequence 17, Application US/10084843  
; Publication No. US20030143243A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; Skeiky, Yasir A.W.  
; Dillon, David C.  
; Campos-Neto, Antonio  
; Houghton, Raymond  
; Vedvick, Thomas S.  
; Twardzik, Daniel R.  
; Lodes, Michael J.  
; Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 355  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/084,843  
; FILING DATE: 25-Feb-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,967  
; FILING DATE: 05-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1872 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-084-843-17

Query Match          98.8%; Score 391.2; DB 12; Length 1872;
Best Local Similarity 99.2%; Pred. No. 1.7e-102;
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1  ACGGCGCGCTCGGATAAATCCAGCTGTCAGGTGGCGAGGATTCGCCATTCGGATC 60
DB   |||||||
QY 758 ACGGCGCGCTCGGATAAATCCAGCTGTCAGGTGGCGAGGATTCGCCATTCGGATC 817
DB   |||||||
QY 61  GGGCAGGCGATCGCATCGCGGCCAGATCCGATCGGTGGGGTCAACCCACCGTTTCA 120
DB   |||||||
QY 818 GGGCAGGCGATCGCATCGCGGCCAAATCCGATCGGTGGGGTCAACCCACCGTTTCA 877
DB   |||||||
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTCGACAAACGCAACGGCGGCACGA 180
DB   |||||||
QY 878 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTCGACAAACGCAACGGCGGCACGA 937
DB   |||||||
QY 181 GTCCAAACGGGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 240
DB   |||||||
QY 938 GTCCAAACGGGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 997
DB   |||||||
QY 241 ATCACCGCGGTGGTCGGGAGCGCTCCGATCAACTCGGCACCGCGATCGGCGCGCTTAAC 300
DB   |||||||
QY 998 ATCACCGCGGTGGTCGGGAGCGCTCCGATCAACTCGGCACCGCGATCGGCGCGCTTAAC 1057
DB   |||||||
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGGCGGCACGGGT 360
DB   |||||||
QY 1058 GGGCATCATCCCGGTGACGTATCTCGGTGAACTGGCAACCAAGTTCGGGCGGCACGGGT 1117
DB   |||||||
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396
DB   |||||||
QY 1118 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 1153
DB   |||||||

RESULT 9
US-10-193-002-17
; Sequence 17, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
;           Skeiky, Yasir A.W.
;           Dillon, Davin C.
;           Campos-Neto, Antonia
;           Houghton, Raymond
;           Vedvick, Thomas S.
;           Twardzik, Daniel R.
;           Lodes, Michael J.
;           Hendrickson, Ronald C.
;
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                   TUBERCULOSIS
;
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
```

```
;
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/09/072,596
;   FILING DATE: 05-MAY-1998
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Maki, David J.
;     REGISTRATION NUMBER: 31,392
;     REFERENCE/DOCKET NUMBER: 210121.417C9
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (206) 622-4900
;     TELEFAX: (206) 682-6031
;
; INFORMATION FOR SEQ ID NO: 17:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1872 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-193-002-17

Query Match          98.8%; Score 391.2; DB 12; Length 1872;
Best Local Similarity 99.2%; Pred. No. 1.7e-102;
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1  ACGGCGCGCTCGGATAAATCCAGCTGTCAGGTGGCGAGGATTCGCCATTCGGATC 60
DB   |||||||
QY 758 ACGGCGCGCTCGGATAAATCCAGCTGTCAGGTGGCGAGGATTCGCCATTCGGATC 817
DB   |||||||
QY 61  GGGCAGGCGATCGCATCGCGGCCAGATCCGATCGGTGGGGTCAACCCACCGTTTCA 120
DB   |||||||
QY 818 GGGCAGGCGATCGCATCGCGGCCAAATCCGATCGGTGGGGTCAACCCACCGTTTCA 877
DB   |||||||
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTCGACAAACGCAACGGCGGCACGA 180
DB   |||||||
QY 878 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTCGACAAACGCAACGGCGGCACGA 937
DB   |||||||
QY 181 GTCCAAACGGGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 240
DB   |||||||
QY 938 GTCCAAACGGGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 997
DB   |||||||
QY 241 ATCACCGCGGTGGTCGGGAGCGCTCCGATCAACTCGGCACCGCGATCGGCGCGCTTAAC 300
DB   |||||||
QY 998 ATCACCGCGGTGGTCGGGAGCGCTCCGATCAACTCGGCACCGCGATCGGCGCGCTTAAC 1057
DB   |||||||
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGGCGGCACGGGT 360
DB   |||||||
QY 1058 GGGCATCATCCCGGTGACGTATCTCGGTGAACTGGCAACCAAGTTCGGGCGGCACGGGT 1117
DB   |||||||
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396
DB   |||||||
QY 1118 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 1153
DB   |||||||
```

```
RESULT 10
US-09-759-143-822
; Sequence 822, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
```

APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09/759,143  
CURRENT FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 934  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-759-143-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.le-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 60  
Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 81  
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCCAACCGTTTCAT 129  
Qy 121 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACACGCGGCGACGA 180  
Db 130 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACACGCGGCGACGA 189  
Qy 181 GTCCAAACGCTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 190 GTCCAAACGCTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249  
Qy 241 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGCGATGGCGGCGCGCTTAAC 300  
Db 250 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGCGATGGCGGCGCGCTTAAC 309  
Qy 301 GGGCATCATCCGCTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGCGGCGACCGGT 360  
Db 310 GGGCATCATCCGCTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGCGGCGACCGGT 369  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 405

## RESULT 11

US-09-780-669-822  
Sequence 822, Application US/09780669  
Patent No. US2002005197A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Jennifer L.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Barrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780.669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-780-669-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.le-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 60  
Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 81  
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCCAACCGTTTCAT 129  
Qy 121 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACACGCGGCGACGA 180  
Db 130 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACACGCGGCGACGA 189  
Qy 181 GTCCAAACGCTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 190 GTCCAAACGCTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249  
Qy 241 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGCGATGGCGGCGCGCTTAAC 300  
Db 250 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGCGATGGCGGCGCGCTTAAC 309  
Qy 301 GGGCATCATCCGCTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGCGGCGACCGGT 360  
Db 310 GGGCATCATCCGCTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGCGGCGACCGGT 369  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 405

## RESULT 12

US-09-822-827-822  
Sequence 822, Application US/09822827  
Patent No. US20020081680A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.534C1  
CURRENT APPLICATION NUMBER: US/09/822.827  
CURRENT FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-822-827-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.le-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 60  
Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 81

QY 61 GGGCAGGCGATGGCGATCCGGGCGCAGATCCGATCGGGTGGGGGTGTCACCCACCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCCGGGCGCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCCCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 180  
Db 130 ATCGGGCCCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 189  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 240  
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 249  
QY 241 ATACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACCGGCTTAAC 300  
Db 250 ATACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACCGGCTTAAC 309  
QY 301 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 360  
Db 310 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 369  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCGCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGACCGCCCGGCC 405

## RESULT 13

US-09-895-793-822

; Sequence 822, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aljun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-822

Query Match 89.4%; Score 354; DB 10; Length 675;

Best Local Similarity 95.7%; Pred. No. 7.1e-92;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGGTCCGATAACTTCCAGTGTCCAGGTGGCGAGGGATTCCGCAATTCGGATC 60  
Db 22 ACGGCGCGGTCCGATAACTTCCAGTGTCCAGGTGGCGAGGGATTCCGCAATTCGGATC 81

QY 61 GGGCAGGCGATGGCGATCCGGGCGCAGATCCGATCGGGTGGGGGTGTCACCCACCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCCGGGCGCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCCCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 180  
Db 130 ATCGGGCCCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 189  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 240  
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 249  
QY 241 ATACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACCGGCTTAAC 300  
Db 250 ATACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACCGGCTTAAC 309  
QY 301 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 360  
Db 310 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 369  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCGCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGACCGCCCGGCC 405

## RESULT 14

US-09-895-814-822

; Sequence 822, Application US/09895814

; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aljun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895,814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-814-822

Query Match 89.4%; Score 354; DB 10; Length 675;

Best Local Similarity 95.7%; Pred. No. 7.1e-92;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGGTCCGATAACTTCCAGTGTCCAGGTGGCGAGGGATTCCGCAATTCGGATC 60  
Db 22 ACGGCGCGGTCCGATAACTTCCAGTGTCCAGGTGGCGAGGGATTCCGCAATTCGGATC 81

Qy	61	GGGAGGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT	120
Db	82	GGGACGGCATGGCGATCGCGGCCAGATC-----AAGCTTCACGACCGTTTCAT	129
Qy	121	ATCGGSCCTACCGCCTTCCTCGGGTTGGGTTTGTCACAACAACGGCAACGGCGCACGA	180
Db	130	ATCGGSCCTACCGCCTTCCTCGGGTTGGGTTTGTCACAACAACGGCAACGGCGCACGA	189
Qy	181	GTCACAACGGTGGTTCGGAGGCGTCCGGCGCAAGTCTCGGCATCTCCACC GGCGACGGT	240
Db	190	GTCACAACGGTGGTTCGGAGGCGTCCGGCGCAAGTCTCGGCATCTCCACC GGCGACGGT	249
Qy	241	ATCACCGCGGTGCAGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGAGCGCGCTTAAC	300
Db	250	ATCACCGCGGTGCAGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGAGCGCGCTTAAC	309
Qy	301	GGGCATCATCCGGTGAGTTCATTCGGTAGCTTGGCAAAACAGTTCGGCGGCACCGGT	360
Db	310	GGGCATCATCCCGGTGAGTTCATTCGGTAGCTTGGCAAAACAGTTCGGCGGCACCGGT	369
Qy	361	ACAGGAAAGTGCACATTGGCCGAGGACCCCGCGCC	396
Db	370	ACAGGAAAGTGCACATTGGCCGAGGACCCCGCGCC	405

RESULT 15

```

US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-678A-822

```

Query Match 89.4%; Score 354; DB 12; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.1e-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCGCGTCCGATAACTTCCAGCTGCCCAGGTGGGCAGGATTCGCCATTCCGATC 60

Db	22	ACGGCCGGTCCGATAACTTCCAGTGTGCCAGGGTGGGCAGGGATTCGCCATTCGGCATC	81
Qy	61	GGGCAGGCGATGGGGATCGCGGGCAGATCGGATCGGCTGGGGGTCAACCCACCGTTTCAT	120
Db	82	GGGCAGGCGATGGCGATCGCGGGCAGATC-----AAGCTTCCCAACCGTTTCAT	129
Qy	121	ATCGGGCCTACCGGCTTCCTCGGCTTGGGTGTGTGCAGACACACGGGCAACGGCGCAGGA	180
Db	130	ATCGGGCCTACCGCTTCTCGGCTTGGGTGTGTGCAGACACACGGGCAACGGCGCAGGA	189
Qy	181	GTCCAACGCGTGTCGGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGGCACGCTG	240
Db	190	GTCCAACGCGTGTCGGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGGCACGCTG	249
Qy	241	ATCACCGGGTCGACGGGGCTCCGATCAACTCGGCCACCGCGATGGCGAGCGCTTAAC	300
Db	250	ATCACCGGGTCGACGGGGCTCCGATCAACTCGGCCACCGCGATGGCGAGCGGCTTAAC	309
Qy	301	GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTTGGCAAAACCAAGTCGGGGGCGCACGGT	360
Db	310	GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTTGGCAAAACCAAGTCGGGGGCGCACGGT	369
Qy	361	ACAGGGACGTGACATTGGCCGAGGACCCCGGGCC	396
Db	370	ACAGGGACGTGACATTGGCCGAGGACCCCGGGCC	405

Search completed: September 5, 2003, 08:54:31

Job time : 1536 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:04 ; Search time 2067 Seconds  
(without alignments)  
4556.300 Million cell updates/sec

Title: US-09-684-215A-3  
Perfect score: 396  
Sequence: 1 acggcggtccgataactt.....tgccggaggagcccccgcc 396

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estin:\*\*  
4: em\_estmu:\*\*  
5: em\_estor:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_htc:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_hc:\*\*  
12: gb\_est3:\*\*  
13: gb\_est4:\*\*  
14: gb\_est5:\*\*  
15: em\_estfun:\*\*  
16: em\_estom:\*\*  
17: em\_gss\_hum:\*\*  
18: em\_gss\_inv:\*\*  
19: em\_gss\_pln:\*\*  
20: em\_gss\_vrt:\*\*  
21: em\_gss\_fun:\*\*  
22: em\_gss\_mam:\*\*  
23: em\_gss\_mus:\*\*  
24: em\_gss\_pro:\*\*  
25: em\_gss\_rod:\*\*  
26: em\_gss\_phg:\*\*  
27: em\_gss\_vri:\*\*  
28: gb\_gss1:\*\*  
29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.2	16.5	289	29	U82114
2	55.6	14.0	603	28	A2934428 BJ_Ba000
3	50.4	12.7	925	29	CNS0091P
4	49.6	12.5	925	29	CNS0091P

5	49.2	12.4	561	28	BH020987
c 6	48.6	12.3	634	10	BE361689
c 7	47.8	12.1	831	29	CC334434
c 8	47.6	12.0	289	29	P864L
c 9	47.4	12.0	935	29	CNS006XK
c 10	47	11.9	1009	29	CNS010EW
c 11	46.4	11.7	515	13	BX424977
c 12	45.8	11.6	472	14	CB620050
c 13	45.6	11.5	515	13	BX424977
c 14	45.4	11.5	459	29	CC349367
c 15	45.4	11.5	830	29	CC331586
c 16	45.2	11.4	538	10	BE215641
c 17	44.8	11.3	502	6	AU191977
c 18	44.6	11.3	839	29	CNS004NB
c 19	44.6	11.3	1201	13	BX456467
c 20	44.6	11.3	1201	13	BX360624
c 21	44.4	11.2	496	12	BJ307565
c 22	44.4	11.2	504	28	BH190076
c 23	44.4	11.2	590	28	BH198840
c 24	44.4	11.2	632	28	BH198865
c 25	44.4	11.2	932	29	CNS00720
c 26	44.2	11.2	844	29	CNS0052P
c 27	44	11.1	493	10	BE355002
c 28	44	11.1	536	10	BG412839
c 29	44	11.1	575	14	CD222185
c 30	44	11.1	587	14	CD229801
c 31	44	11.1	618	14	CD221131
c 32	44	11.1	644	14	CD226234
c 33	43.8	11.1	527	14	CB636890
c 34	43.8	11.1	538	14	CB636749
c 35	43.8	11.1	767	14	CB676645
c 36	42.6	10.8	442	12	BM523115
c 37	42.6	10.8	498	10	BG411096
c 38	42.4	10.7	638	14	CD220173
c 39	42.4	10.7	959	29	BZ549048
c 40	42.2	10.7	578	14	CB636482
c 41	42.2	10.7	1201	13	BX360624
c 42	42.2	10.7	1201	13	BX460099
c 43	42	10.6	755	10	BE705003
c 44	42	10.6	1004	13	BU772046
c 45	42	10.6	1033	29	BZ561390

#### ALIGNMENTS

RESULT 1  
U82114  
LOCUS  
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.  
ACCESSION U82114  
VERSION U82114.1 GI:3647212  
KEYWORDS GSS.  
SOURCE Mycobacterium leprae  
ORGANISM Mycobacterium leprae  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 289)  
AUTHORS Silbag,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.  
TITLE Characterization of a 34-kilodalton protein of Mycobacterium leprae that is homologous to the immunodominant 34-kilodalton antigen of Mycobacterium paratuberculosis  
JOURNAL Infect. Immun. 66 (11), 5576-5579 (1998)  
MEDLINE 99003163  
PUBMED 9784577  
COMMENT Contact: Silbag FS  
Microbiology  
Colorado State University  
Fort Collins, CO 80523, USA  
Egilmiet,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use of an ordered cosmid library to deduce the genomic organization of Mycobacterium leprae. Mol. Microbiol. 7 (2), 197-206 (1993)

289 bp DNA linear GSS 16-FEB-2001  
Mycobacterium leprae genomic clone

Query Match	14.0%	Score 55.6;	DB 28;	Length 603;
Best Local Similarity	54.4%;	Pred. No. 0.00099;		
Matches	112;	Conservative 0;	Mismatches 94;	Indels 0; Gaps 0;
QY	140	TCGGCTTGGGTGTTGTCACAACACGCGCAACGCGCGACGAGTCCACACGCGTGGTCGGGA	199	
Db	363	TCGCGGACAGCCCTCGGCATGAAGAAGGCCGAAGGGCGCTGGTGGCGAGACCCGACGCGA	422	
QY	200	CGCTTCGGCGGGCAAGTCTCGGCATCTCCACCGCGGACGCTGATCACCGCGTTCGACGGCG	259	
Db	423	ACGGTTCGGCGGGCAAGCGCGGATCGAGTCCGGGACGCTGATCACCTCGGTCAACGGCG	482	
QY	260	CTCCGATCAACTCGGCCACCGCGATGCGCGACGCGCTTAACGGGCAATCATCCCGGTGACG	319	
Db	483	AATCCGTCGAAGGACGCCCGGAGCTCGCCGCGACCATCGGCGGCATGGCGCCGTCGCGA	542	
QY	320	TCATCTCGGTGACCTGCGCAACCAAG	345	
Db	543	TCGTGAAGCTTAACGTCTGCACAAG	568	
RESULT 3				
CNS0091P				
LOCUS	CNS0091P	925 bp	DNA	linear
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL053013			
VERSION	AL053013.1 GI:4934461			
KEYWORDS	GSS.			
SOURCE	Drosophila melanogaster (fruit fly)			
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 925)			
REFERENCE	Genoscope.			
AUTHORS	Direct Submission			
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequences :			
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp; the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .			
FEATURES	Location/Qualifiers			
source	1..925			
	/organism="Drosophila melanogaster"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:7227"			
	/clone_lib="BACR19D16"			
	/clone_lib="RPCI-98"			
	/note="end : TET3"			
BASE COUNT	120 a	61 c	61 g	172 t
ORIGIN	511 others			
Query Match	12.7%	Score 50.4;	DB 29;	Length 925;
Best Local Similarity	15.4%;	Pred. No. 0.02;		
Matches	53;	Conservative 158;	Mismatches 133;	Indels 1; Gaps 1;
QY	53	TTCCGATCGGGAGCGGATCGCGCGGCGAGATCCGATCGCGTGGGGTCAACCA	112	



```

Db      553 TTSSGSGYGGKSGSGBSCSCSSCSCSSSCSCBCCSCSSSYCCSSSSSSKCSS 612
QY      113 CCGTTATATACGGCCCTACCGCTTCCTCGCTTGGGTGTTCGACAAACAGGCAACG 172
Db      613 TSBSGCCSCCSCSVCVGTSCSSSSCSCSSSSSTSSSTSSKSSSGSSSSSYTTS 672
QY      173 GCGACAGTCCACACGGG-TGGTGGGAGCGCTCCGCGGCAAGTCTTCGCATCTCCACC 231
Db      673 KTSASGSGSWAGGSGTSTSSSSSSTSTSSSSVSSTSTSSSSVSSTSSSSGSS 732
QY      232 GCGACGTATACCGCGGTTCGACGGCTCCGATCAACTCGCCACCCTGGGATGGCGAC 291
Db      733 SSSTSSBBSSTSSSSSSSYSSSTCCCTCCSSSYSSSTSSSTSSSTSSGSSSV 792
QY      292 GCGCTTAACGGGATCATCCGCTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGC 351
Db      793 GTSSSDSTTCCSCCCTCCCTTCCTYBMCYTSTSCGSSSSSGKGGVTKCGCGCGSS 852
QY      352 GGCACGGTACAGGAACGTGACATTCGGCGAGGACCCCGCGCC 396
Db      853 TNGMBGTSSACSSSSSCSSSSSSSSKSSASSSSVSSSGSGVS 897

RESULT 4
CNS0091P/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN
Query Match 12.5% Score 49.6; DB 29; Length 925;
Best Local Similarity 12.8% Pred. No. 0.031;
Matches 43; Conservative 157; Mismatches 136; Indels 0; Gaps 0;

```

```

QY      60 CCGGACGCGATCGCGATCGCGGCGCAGATCCGATCGGTGGGGGTTCACCCACCGTTCA 119
Db      900 SNSSBSCSSSSSSSTSSSSSSBSSSSSSSGTSSACVKCNASSSGCGCGMA 841
QY      120 TATCGGCGCTACCGCTTCTCGCTTGGTGTTCGACAAACGCGACGCGCGACG 179
Db      840 BCMCMSSSSSCCGSASARGVKVRASGAGKRGSGGASGASHSSSSACBSSSSSCSASCW 781
QY      180 AGTCCAAACGGTGTTCGGGAGCGCTCCGCGCGCAAGTCTCGGATCTCCACCGCGACGT 239
Db      780 SASSSSSASRRSRSGGAGGSGASSRSSSSSSSSASAGSVSSASSSSSSSSVSCSS 721
QY      240 GATCACCGCGCTACGCGCTCCGATCAACTCGGCACCGCGATGCGCGACGCGCTTAA 299
Db      720 VASSMCSBSSSSASASSSSSSSSSASCSCCCTSWSCSTSSASMSAARSSSSSS 661
QY      300 CGGCGATCATCCCGGTGACGTCTCTCGGTGACTGCGCAACCAAGTCGGCGCGCACGG 359
Db      660 CSSSSMSASSSSSSSSSSSSSSSGSAGCBMSMSSGGSGSVSSASGMSSSSVSSG 601
QY      360 TACAGGGAACGTGACATTCGCGGAGGACCCCGCGC 395
Db      600 GRSSGSGGGVGGSGSSSGSGSGSVSCSS 565

RESULT 5
BH020987
LOCUS
DEFINITION
P864c.d_LEISHPAC-left.1 Leishmania major Friedlin Cosmid Genomic
Library Leishmania major genomic clone P864c, genomic survey
sequence.
ACCESSION
BH020987
VERSION
BH020987.1 GI:14202102
KEYWORDS
GSS.
SOURCE
Leishmania major
ORGANISM
Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.
REFERENCE
1 (bases 1 to 561)
Mylers, P.J., Vogt, C., Cawthra, J., Klacking, M., Marty, A., Mack, J.,
Munden, H., Nguyen, D., Robertson, L., Sisk, E., Fazelinia, G., Aggarwal
, G., Nelson, S., Seyler, A., Worthey, E. and Stuart, K.
Leishmania major Friedlin Cosmid End Sequences
Unpublished
Other GSSs: P864c.d_LEISHPAC-right.1
Contact: Myler PJ
Seattle Biomedical Research Institute
4 Nickerson Street, Seattle, WA 98109-1651, USA
Tel: 206 284-8846
Fax: 206 284-0313
Email: mylerpj@sbri.org
Seq primer: LEISHPAC-Left
Class: PAC end.
FEATURES
Location/Qualifiers
1..561
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="P864c"
/lab_host="E. coli ED8767"
/clone_lib="Leishmania major Friedlin Cosmid Genomic
Library"
/note="Vector: cLHYG; Site 1: BamHI; Genomic DNA from
Leishmania major Friedlin was partially digested with
Sau3AI, size selected, and ligated with BamHI-digested
cLHYG cosmid vector DNA. 9216 clones were picked and
arrayed. Library construction is described in Ivins et
al., Genomics Research, 8:135-145 (1998). The cLHYG
vector (Acc. No. CVU59231) is described in Ryan et al.,
Gene, 131:145-150 (1993)"
BASE COUNT 108 a 173 c 182 g 98 t
ORIGIN

```

Db	268	ACGACGCTGTCGTCGCTGCTCTCTCTCGGAGTCGGGCGCGGAAAGCCCTCTCTGCTGCTGCT	209
Qy	197	GGAGCTCTCGGCGGGCAAGTCTCGGCATCTCCACCGCGCAGTGCATCACCGCGTGGACG	256
Db	208	GCGCCGCGCGCGGGCGGCGCTACTGCTCCGGCGTGGCGGCTTGGCGTGGCTGGACTTGT	149
Qy	257	GCCTCCGATGAACCTCGGCCACCGCATGCGGCAGCGCTTAAAGGGCATCATCCCGGTG	316
Db	148	GGCGCGCGAGCGCTGTACGACGCGAAGGCCCTCCGACACCGCGCACCGGAACCGGA	89
Qy	317	ACGTCACTCTGGTGACCTG	335
Db	88	GCTGCTCCTCGTGGCGCAG	70

RESULT 7	
CC334434/c	
LOCUS	831 bp DNA linear GSS 16-MAY-2003
DEFINITION	CGCAG83TV ZM_0.7.1.5_KB Zea mays genomic clone ZMMBMa0365N21, genomic survey sequence.
ACCESSION	CC334434
VERSION	CC334434.1 GI:30803847
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
AUTHORS	1 (bases 1 to 831) Whitelaw,C.A., Fourckenbush,J., Van Aken,S., Utterback,T., Resnick ,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek ,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics
TITLE	Unpublished
JOURNAL	Contact: Cathy Whitelaw
COMMENT	TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.

```

location/mutational/cis
1. .83%
/organism="zebra mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/cloned_lib="ZM0365N21"
/cloned_lib="ZM_0.7-1.5_KB"
/note=vector: pBCSK+; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

BASE COUNT      149 a    280 c     290 g   112 t
ORIGIN
Query Match          12.1%; Score 47.8; DB 29; Length 831;
Best Local Similarity 57.1%; Pred. No. 0.086;
Matches 109; Conservative 0; Mismatches 77; Indels 5; Gaps 1;

Qy       164 ACGGCAACGCGCCTAGATCCAACGGGTGTCGGGAGCGCTCGGGGGGAAGTCTCGGCA 223
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       777 ACGGGCACGCCGCAGAAGGCGCGTTTGCTCGCGGAGCATGGCGG-----CTGCGCG 723

Qy       224 TCTCCAACGCGGAGGTGATACCACGCGGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGA 283
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       722 TCGCCACCTCTGAGGTGCGCGCGCTCGACGCCGCCGCCGCTGACAGCGGCCGCCGCCG 663

Qy       284 TGCGGACGCGCTTAACGGGCGATCATCCCCGGTGACGTCACTCGGTGACCTGGCAAACCA 343
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       662 GC GGCTTCCGGTCCGGGGCGGCGCTCTCGCGGACGTCTCTCTCTCTCTCTCTCTCGCGTCG 603

Qy       344 AGTCCGGCGGC 352
||||| |||||

```

```

Db      602 CCCC GCCCGC 592

RESULT 8
P864L
LOCUS   Leishmania major Friedlin PAC P864 left end-sequence, genomic
DEFINITION
ACCESSION AL390548
VERSION   AL390548.1 GI:9501524
KEYWORDS GSS
SOURCE   Leishmania major
ORGANISM Leishmania major
          Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
REFERENCE 1 (bases 1 to 289)
AUTHORS  Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and
          Smith,D.F.
          A physical map of the Leishmania major Friedlin genome
          Genome Res. 8 (2), 135-145 (1998)
          98146435
          9477341
          2 (bases 1 to 289)
          Taylor,R.G., Huckle,E.E.J., Ivens,A.C., Rajandream,M.A. and
          Barrell,B.G.
          Direct Submission
          Submitted (24-JUL-2000) Leishmania major Friedlin genome sequencing
          project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
          Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
          alicat@sanger.ac.uk
          see http://www.ebi.ac.uk/parasites/leish.html
          Details of Leishmania sequencing at the Sanger Centre are available
          at http://www.sanger.ac.uk/projects/Lmajor/
          The primer sequence can be obtained from alicat@sanger.ac.uk.

FEATURES
    source
        1..289
            /organism="Leishmania major"
            /mol_type="genomic DNA"
            /strain="Friedlin"
            /db_xref="taxon:5664"
            /clone="PAC P864"
BASE COUNT 49 a 86 c 108 g 46 t
ORIGIN
Query Match 12.0%; Score 47.6; DB 29; Length 289;
Best Local Similarity 62.7%; Pred. NO. 0.084;
Matches 74; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 159 CAACACGGGCAACGGCGACGATCCACGCGTGTGCGGGAGCGTCCGGCGGCAAGTCT 218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24 CGAGAGCGGCATCCCGCGCCCTGTAGTCGAGGGCGCGTCCGCGTCCCGGGTCCGGT 83
QY 219 CGGCATCTCCACCGCGACGTGATCACCGCGGTGACGCGCGTCCGATCAACTCGGCC 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 CGGATTCACCCGGCGACGTGATCTCGCCTTCAACGGCAGCGGATCAAGTCAAGTC 141

RESULT 9
NS006XK
LOCUS   Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BAC14N09 of RPCI-98 library from Drosophila melanogaster (fruit
ACCESSION AL066051
VERSION   AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS  Genoscope.

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
ECORI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
    source
        1..935
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BAC14N09"
            /clone_lib="RPCI-98"
            /note="end : T7"
BASE COUNT 257 a 170 c 162 g 96 t 250 others
ORIGIN
Query Match 12.0%; Score 47.4; DB 29; Length 935;
Best Local Similarity 28.2%; Pred. NO. 0.11;
Matches 111; Conservative 89; Mismatches 193; Indels 0; Gaps 0;
QY 4 GCCCGCTCCGATACTCCAGCTGTCCAGGTGGGAGGATTCGCCATTCCGATCGG 63
    ||:| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 523 GCSTCCMYMSSSVSCSCSGTCCGTCCTCKMCSCTYGGCKGCGCCSTCSSSSSCS 582
QY 64 CAGCGCATGGCGATCGGGGCCAGATCGGATGGGGGTGACCCACCGTTTCATATC 123
    : : : : |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 583 BBSYSTCCTBCTKSSGCSGTSGTCGGGGSGCGGCGCGCGGCGGCGSS 642
QY 124 GGGCTTACCGCTTCTCGGCTTGGTGTTCGACACACAGGCAAGCGCGCAGAGTC 183
    ||::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 643 GSSGSSGSSGSSGSSGSSGSSGSGGSGGCGCGCGCGCGCGCGSGCGSG 702
QY 184 CAACGCTGTGTCGGAGCGCTCCGCGCAAGTCTCGGCATCTCCACGCGCAGTGC 243
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 703 CGSGCGSGGGCGCGCGCGCGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 762
QY 244 ACCGCGTTCGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAACGG 303
    : : : : |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 763 MSMCGGGGSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGSG 822
QY 304 CATCATCCCGTGTAGTTCATCTCGGTACCTGGCAACCAAGTCGGGCGGACGGGTACA 363
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 823 GSCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGSGSS 882
QY 364 GGAACGCTGACATTGGCGGAGGACCCCGCGCC 396
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 883 GSGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGSS 915

RESULT 10
NS010EW/C
LOCUS   Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN03P19 of DrosBAC library from Drosophila melanogaster (fruit
ACCESSION AL098832
VERSION   AL098832.1 GI:5610493
KEYWORDS GSS.
SOURCE   Drosophila melanogaster (fruit fly)

```

CB620050 472 bp mRNA linear EST 08-APR-2003  
OSIIIEa04P23.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA  
clone OSIIIEa04P23 5', mRNA sequence.  
CB620050  
CB620050.1 GI:29615037  
EST.  
Oryza sativa (indica cultivar-group)  
Oryza sativa (indica cultivar-group)  
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae;  
Ehrhartoideae: Oryzeae; Oryza.  
1 (bases 1 to 472)  
Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona

1000

THE UNIVERSITY OF CHICAGO

```

/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0387M02"
/clone_lib="ZM_0_7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
BASE COUNT      62 a 175 c 187 g 35 t
ORIGIN
Query Match      11.5%; Score 45.4; DB 29; Length 459;
Best Local Similarity 57.5%; Pred. No. 0.31;
Matches 103; Conservative 0; Mismatches 71; Indels 5; Gaps 1;

QY 164  ACGGCAACGGCGGACGAGTCCAAAGCGGTGTGCGGGAGCGCTCGGGCGGCGAAGTCTCGGCA 223
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 446  ACGGCGACGGCGGAGAAAGCGCGTTCGTCTGTCGGGAGCATGCGCGG-----CTGCCCG 392
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 224  TCTCCACGGCGGACGTGATCACGCGGTCTGACGGCGCTCCGATCAACTTCGGCCACCGCA 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391  TCGCCACCTTCGAGTGTGCGGCGCTGACGCCGCCGCCGCGCTGCACGGCGGCGCGGC 332
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 284  TGGCGGACGGCGTTTACGGGATCATCCCGGTGACGTATCTCGGTGACCTTGGCAACC 342
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331  GCAGTTTTCGGGTCCGGCGGCGCTCTCCGCGACGTCCTCTCCTCTCGCGCGTCC 273
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 15	CC331586	830 bp	DNA	linear	GSS 16-MAY-2003
LOCUS	OCGBH15TV ZM.0_7.1_5_KB	zea mays	genomic clone	ZMMBMA0344C06,	
DEFINITION	genomic survey sequence.				
ACCESSION	CC331586				
VERSION	CC331586.1	GI:30800757			
KEYWORDS	GSS.				
SOURCE	zea mays				
ORGANISM	zea mays				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	1 (bases 1 to 830) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick ,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek ,R.W., Nunberg,A., Robbins,D. and Lakey,N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished				
COMMENT	Contact: Cathy Whitelaw TIGR				
	9712 Medical Center Drive, Rockville, MD 20850, USA				
	TEL: 301-838-5843				
	Fax: 301-838-0208				
	Email: whitelaw@tigr.org				

```

Class: smeared ends.
Location/Qualifiers
1. .830
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBma0344C06"
  /clone_lib="2M_0.7_1.5_KB"
  /note="vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
  methylation filtered genomic DNA library"
134 a      244 c      337 g      115 t
BASE COUNT
ORIGIN

```

Query Match	11.5%	Score 45.4	DB 29	Length 830
Best Local Similarity	57.5%	Pred. No. 0.33		
Matches 103	Conservative	0	Mismatches 71	Indels 5
Gaps 1				
Qy	164	ACGGCAACGGCGACAGTCCACGGCTGTCTCGGAGCGTCCGGCGGCAAGTCTCGGCA	223	

三

```

389  ACGGGCACGGGCACAGAGCCCGCTTCGTGTCCTCCCGCAGCATCGGGG-----CTGCCCCG 335
224  TCTCCACGGGAGCTGATCACCGGGTGCAGCGGGCTCCGATCAACTCGGCACACGGCA 283
334  TCGCCACCTTGAGTGGCGGGCTGCAGCGCGCCCGCGCTGCACGGCGCGCGCGC 275
284  TGGCGGACGCGCTTAACGGGATCATCCGGTGCAGTTCATCTCGTGCCTGGCAAAAC 342
274  GCGCTTCGGGTCCGGCGGGCGCTTCCTCCGGAGCTCCTCTCCTCTCTCTCTCTCTCT 216

Search completed: September 5, 2003, 08:27:15
Job time : 2070 secs

```

Search completed: September 5, 2003, 08:27:15  
Job time : 2070 secs

[illegible]

RESULT 15	
CC331586/c	
LOCUS	830 bp
DEFINITION	linear DNA
	OG0BH15TV 2M_0.7_1.5-KB Zea mays genomic clone ZMMBMA0344C06,
	genomic survey sequence.
ACCESSION	CC331586
VERSION	CC331586.1
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 830)
AUTHORS	Whitlaw,C.A.; Quackenbush,J.; Van Aken,S.; Utterback,T.; Resnick A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE	Consortium for Maize Genomics
JOURNAL	Unpublished
COMMENT	Contact: Cathy Whitlaw

```

CONTACT: Cathy Whiteley
TYGR Medical Center Drive, Rockville, MD 20850, USA
TEL: 301-838-5843
FAX: 301-838-0208
Email: whitelawetgr.Org
Seq primer: TF
Class: sheared ends.
FEATURES
    source
        1..830
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBMA0344C06"
            /note="vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                    methylation filtered genomic DNA library"
134 a      244 c      337 q      115 t
BASE COUNT

```

Query Match	11.5%	Score 45.4	DB 29	Length 830
Best Local Similarity	57.5%	Pred. No. 0.33		
Matches 103	Conservative	0	Mismatches 71	Indels 5
Gaps 1				
Qy	164	ACGGCAACGGCGACAGTCCACGGCTGTCTCGGAGCGTCCGGCGGCAAGTCTCGGCA	223	

三

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:52:39 ; Search time 2356.88 Seconds  
(without alignments)  
2291.192 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGGQGAIPA.....QTKSGGTRGNTVLAEGPPA 132

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US09684215/runat\_05092003\_072202\_559/app\_query.fasta\_1.853  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215.ecgn\_1.1.3962 @runat\_05092003\_072202\_559 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*

- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pin:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	675	100.0	1068	6	AX005788 Sequence
2	675	100.0	1143	6	AX005790 Sequence
3	675	100.0	2287	6	AR303127 Sequence
4	675	100.0	11700	1	MTC1418B Mycobacteri
5	675	100.0	14029	1	AE006925 Mycobacte
6	675	100.0	343050	1	BX248334 Mycobacte
7	670	99.3	447	6	AR169152 Sequence
8	670	99.3	447	6	AR182442 Sequence
9	670	99.3	447	6	AR194825 Sequence
10	670	99.3	447	6	AR233097 Sequence
11	670	99.3	447	6	AX429596 Sequence
12	670	99.3	447	6	BD006325 Compounds
13	670	99.3	447	6	BD006445 Compounds
14	670	99.3	447	6	BD069285 Compounds
15	670	99.3	1872	6	AR169165 Sequence
16	670	99.3	1872	6	AR182455 Sequence
17	670	99.3	1872	6	AR194838 Sequence
18	670	99.3	1872	6	AR233110 Sequence
19	670	99.3	1872	6	AX429609 Sequence
20	670	99.3	1872	6	BD006338 Compounds
21	670	99.3	1872	6	BD006458 Compounds
22	670	99.3	1872	6	BD069298 Compounds
23	632	93.6	675	6	AX201049 Sequence
24	632	93.6	675	6	AX267848 Sequence
25	632	93.6	822	6	AR277645 Sequence
26	632	93.6	822	6	AR277645 Sequence
27	632	93.6	822	6	AX369152 Sequence
28	632	93.6	894	6	AX351489 Sequence
29	632	93.6	900	6	AR220690 Sequence
30	632	93.6	900	6	AX365960 Sequence
31	632	93.6	915	6	AR261281 Sequence
32	632	93.6	915	6	AX201061 Sequence
33	632	93.6	915	6	AX267860 Sequence
34	632	93.6	945	6	AR277644 Sequence
35	632	93.6	945	6	AX369151 Sequence
36	632	93.6	1012	6	AR220689 Sequence
37	632	93.6	1012	6	AX365958 Sequence
38	632	93.6	1155	6	AX369165 Sequence
39	632	93.6	1203	6	AX201078 Sequence
40	632	93.6	1203	6	AX267877 Sequence
41	632	93.6	1464	6	AR229410 Sequence
42	632	93.6	1464	6	AX156105 Sequence
43	632	93.6	1464	6	AX361955 Sequence
44	632	93.6	1557	6	AR229398 Sequence
45	632	93.6	1557	6	AX156089 Sequence

ALIGNMENTS

RESULT 1

```
AX005788
LOCUS AX005788 1068 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 907 from Patent WO9909186.
ACCESSION AX005788
VERSION AX005788.1 GI:9928795
KEYWORDS Mycobacterium tuberculosis
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
AUTHORS Portnoi, D. and Guigueno, A.
TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors
comprising same and uses for diagnosing and preventing tuberculosis
JOURNAL Patent: WO 9909186-A 907 25-FEB-1999;
PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES
source 1..1068
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"
1..1068
/contig="50D"
/codon_start=1
/transl_table=11
/protein_id="CAC05178.1"
/db_xref="GI:9928796"
/translation="MSNSRRSLRWSLLVLAAGLGLATAPAAQAPALSQDRFAD
FPALPLDPSAMVAQVGPQVAVNTKLGYNNAVGAGTGVIDPENGVLNNHVIAGATD
INAFSVGGQTYGVVDGTRDQVAVIQLRAGGLPSAAIIGGVAVGEPVAVMNSG
GOGGTPRAVPQVAVLQGTQVQSDSLTGAETLNLQIDFAAIQPDGSGGPVNLGQ
VNGMTAASDNFQLSQGGQFAIPIGQAMATAGQIRSGGSPTHVIGPTAFGLGLVVD
NNGCARVQRVVSAPAAASLGISTGDTVITAVDGAIPNSATAMADALNGHHFPGDVISVT
WQTKSGGTRTGNVTLAEGPPA"
BASE COUNT 169 a 352 c 375 g 172 t
ORIGIN
Alignment Scores:
Pred. No.: 3.25e-42 Length: 1068
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AX005788 (1-1068)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db ACAGCGCGGCTCCGATTAACCTCCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 729
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db GGGCAGGCGATGCCATCCGGCGCCAGATCCATCGGTGGGGTCCACCCACCTTCAT 789
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAlaArg 60
Db ATCGGCGCTACCGCTCTCTCGCTTGGGTGTGGTGTGTCGACAAACAGCGACGACGA 849
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrClyAspVal 80
Db GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGCGGACGATG 909
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db ATCACCAGCGGTGCGAGCGCTCCGATCAACTCGGCCACCGGATCGCGCGGCGCTTAAC 969
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACCAAGTCGGCGGCGACGCT 1029
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
```

```
|||||
DB 1030 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 1065
|||||
RESULT 2
AX005790 LOCUS AX005790 1143 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 909 from Patent WO9909186.
ACCESSION AX005790
VERSION AX005790.1 GI:9928797
KEYWORDS Mycobacterium tuberculosis
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
AUTHORS Portnoi, D. and Guigueno, A.
TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors
comprising same and uses for diagnosing and preventing tuberculosis
JOURNAL Patent: WO 9909186-A 909 25-FEB-1999;
PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES
source 1..1143
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"
<4..1143
/codon_start=1
/transl_table=11
/protein_id="CAC05179.1"
/db_xref="GI:9928798"
/translation="KNPAARTLKAGTISDSYPTDGTGGYGMNSRRSLRWSLLVLA
VGLGLATAQAAPALSDRFADFPALPLDPSAMVAQVGPVAVNTKLGYNNAVGA
GTGVIDPENGVLNNHVIAGATINAFSVGGQTYGVVDGTRDQVAVIQLRAG
GLPSAAIIGGVAVGEPVAVMNSGQGTTPRAVPQVAVLQGTQVQSDSLTGAETL
GLIQDFAAIQPDGSGGPVNLGQVNGMTAASDNFQLSQGGQFAIPIGQAMATAG
IRSGGSPTHVIGPTAFGLGLVVDNNGCARVQRVVSAPAAASLGISTGDTVITAVD
PINSATAMADALNGHHFPGDVISVTWQTKSGGTRTGNVTLAEGPPA"
BASE COUNT 189 a 373 c 395 g 186 t
ORIGIN
Alignment Scores:
Pred. No.: 3.48e-42 Length: 1143
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AX005790 (1-1143)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db ACAGCGCGGCTCCGATTAACCTCCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 804
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db GGGCAGGCGATGCCATCCGGCGCCAGATCCATCGGTGGGGTCCACCCACCTTCAT 864
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAlaArg 60
Db GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGCGGACGATG 924
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrClyAspVal 80
Db GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGCGGACGATG 984
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db ATCACCAGCGGTGCGAGCGCTCCGATCAACTCGGCCACCGGATCGCGCGGCGCTTAAC 1044
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
```



```

|||||
1045 GGGCATCATCCGGTACGTCATCTCGTGCACCTGGCAACCAAGTGGCGCGCACGCGT 1104
|||||
121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
1105 ACAGGGAACGTGACATGGCCGAGGAGACCCCGGCC 1140
|||||

RESULT 3
AR303127
LOCUS AR303127 2287 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544522.
ACCESSION AR303127
VERSION AR303127.1 GI:31691855
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2287)
AUTHORS Skeiky,Y., Alderson,M. and Campos-Reto,A.
TITLE Fusion proteins of mycobacterium tuberculosis antigens and their
uses
JOURNAL Patent: US 6544522-A 1 08-APR-2003;
FEATURES
source Location/Qualifiers
BASE COUNT 372 a 717 c 814 g 381 t 3 others
ORIGIN
1..2287
/organism="unknown"

Alignment Scores:
Pred. No.: 6,93e-42 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x AR303127 (1-2287)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 63 ACGCGCGCTCCGATACCTTCCAGCTGTCCAGGGTGGGAGGATCGGCATTCGATC 122
|||||
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
|||||
Db 123 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCACCGTTCAT 182
|||||
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60
|||||
Db 183 ATCGGGCCTACCGCCTTCCTCGGCTGGGTGTTCGACAAACACGCGACGCGCACGA 242
|||||
Qy 61 ValClnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
|||||
Db 243 GTCCACGCGTGTTCGGGAGCGTCTCGCGCGCAAGTCTCGGCATCTCCACCGCGCGTG 302
|||||
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
|||||
Db 303 ATCACCAGCGTTCGACGCGCTCCGATCAATCGGCCACCGCATGGCGGACGCGCTTAAC 362
|||||
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpClnThrIlySerGlyGlyThrArg 120
|||||
Db 363 GGGCATCATCCGGTACGTCATCTCGGTACCTGGCAACCAAGTGGCGCGCACGCGT 422
|||||
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 423 ACAGGGAACGTGACATGGCCGAGGAGACCCCGGCC 458
|||||

RESULT 4
MTCI418B
LOCUS MTCI418B 11700 bp DNA linear BCT 02-SEP-2002
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 7/162.
ACCESSION Z96071 AL123456
VERSION Z96071.1 GI:3242254
KEYWORDS

```

```

SOURCE Mycobacterium tuberculosis H37Rv
ORGANISM Mycobacterium tuberculosis H37Rv
Bacteria: Actinobacteria: Actinobacteridae: Actinomycetales:
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagsis,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skellton,S.,
Squares,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and
Barrell,E.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)
98295987
9634230
2 (bases 1 to 11700)
Parkhill,J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 20, 1998 this sequence version replaced gi:2181960.
Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TParse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.
FEATURES
source Location/Qualifiers
1..11700
/organism="Mycobacterium tuberculosis H37Rv"
/mol_type="genomic DNA"
/db_xref="H37Rv"
/db_xref="taxon:83332"
1309..11293
/organism="Mycobacterium tuberculosis H37Rv"
/mol_type="genomic DNA"
/strain="H37Rv"
/db_xref="taxon:83332"
/cloned="I418B"
10774..>11700
/organism="Mycobacterium tuberculosis H37Rv"
/mol_type="genomic DNA"
/strain="H37Rv"
/db_xref="taxon:83332"
/cloned="CI5"
<1..1409
/note="fragment designated v031. Does not represent a
physical clone"
83..88
/note="possible RBS, AAGGAG, for RV0119"
97..1674
/gene="fadD7"
97..1674
/gene="fadD7"
/note="fadD7"
/note="RV0119, (MTV031.13-MTCI418B.01), len: 525.fadD7,

```

PossibleCoenzyme A-ligase similar to 4-coumarate:CoA ligase of many organisms e.g. PTU39405\_1 (537 aa). Contains PS00455putative AMP-binding domain signature. FASTA scores: gpiU39405|ptu39405\_1 Pinus taeda xylem 4-coumarate:CoA (537 aa) opt:483z-score: 526.1 E(); 8.3e-22; 28.2% identity in 440 aaoverlap score is 0.896"

```

/codon_start=1
/transl_table=11
/product="fadD7"
/db_xref="GI:3242255"
/db_xref="SPTREMBL:O07169"
/translation="MASDFGRIADLVVEAATRLPEAPALVVTADRIAISHRDRLARLV
DELQAQTRSLGPDGRVALRMGSNAEFVALLAASRADLVVPLDLPALPTEQQRVS
QAAGARVYLIDADGPHDRAEPTTRWPLTVNVGDSGSPSGTSLVHLDAAFPVNPATS
TPEGLRPDDAMITGGTGTGLPKVWPMTNIASSVRAITGYRLSPDRAFAVMPPLY
HGHLIASLALATAGGAVSLPARGFSAHFWDDIKAVGATVTAFTTHOILLERS
ATPESGRKPAALFIRSCSAPLTAQALALQTEFAAPVCAFGMTATHQVTTQIEG
IDQETPVSTGLVGRSTGAQIRIVGSDGLPLPAGVGEIWLRTGTVVRYGLGDPIT
AANFTDGLWRTGDLGSLSAAGDLSIRGRIKELINRGGEKISPERVEGVLASHPNVMEA
AVFGVPHQLYGEAAVIVPRESAPPTREELVQCRERLAAFEIPASFOEASGLPHTA
KGLSDRRVAERFGHSV"
610..645
/misc_feature
/gene="fadD7"
/notes="PS00455 Putative AMP-binding domain signature,
[LIYWFY].[2][STG][STAG]G[ST][STEE][SG].[PASLIVM][KR], info
count = 22.0"
/gene="fusA2"
/complement(1675..3819)
/complement(1675..3819)
/gene="fusA2"
/notes="Rv0120c, (MTC1418B.02c), fusA2, len: 714.
Elongation factor g, similar to eg. EFG_ECOLI P02996
Elongation factor g (ef-g). Escherichia coli (703 aa),
fasta scores: opt:1049 z-score: 1105.5 E(): 0..32.5%
identity in 717 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Also similar to M.
tuberculosis fusA, MTCY210.01 (39.1% identity in 299 aa
overlap). Similar to spIP30767IEFG_MYCLE_ELONGATION FACTOR
G (EF-G). (701 aa); 31.7% identity in 710 aa overlap."
/codon_start=1
/transl_table=11
/product="fusA2"
/protein_id="CAB09448.1"
/db_xref="GI:2181962"
/db_xref="SWISS-PROT:O07170"
/translation="MADRNASOGAAAAPTANGPGGVNRVNVLVGSPGGKTLIEALL
VAKVLSPGSGVTGTTVCDPDEAIEIQORSVGLAVASLAYDGIKVLNVDTPGVADEV
GELAGLRADCALEFVTAANEGVDEPTKSLWQECQVGMPPRAVITKLDHARANYREA
LTAQDAFGRKVLPLYPSGDGLIGLSQALYEDAGKRITRTPAESDTERIEEARGA
LIEGIIIESEDESLEMYLGGETIDESVLIDLEKAVARGSFVPIVPCSTGVGTILE
LLEVATRGPSPMEHPLPEVFTPGQVPHAEACNDADAPLLAEVYKVTSDPYVGRVSLV
RVFSGTIRPDVTTHVSGHFSFFGGGTSTHPDHDEDERIGVLSPFLGKQORPAAVYV
AGDCAIGKLSRAETGDTLSKAEPLVLKPMTPPELLPIAIAHAHATDEDKLSVGLG
RLAADFTPLREQNQETHQVYWCNMGHAHAGVLDLTANRYGVSDIIEPLVRETF
AGNAGGHRH1KQSGHGQYGVCDIEVEPLPEGSGFEFLDKVGVGAVPRQIPNVEKG
VRAQMDKGVHAGYVPVDIRVTLTDGKAHVSDDSAFQFQAGALALREAAATKIVILLE
PIDETSVLPDDFDFGVGRIGLSSRRRVLGTPETAGHDRTVIKAEVPPQVELTRYAIDLR
SLANGAASFTSFARYPEMPESAARVAKAGAG"
complement(3709..3732)
/gene="fusA2"
/notes="PS00017 ATP/GTP-binding site motif A (P-loop)"
/complement(3956..4390)
/gene="Rv0121c"
/complement(3956..4390)
/gene="Rv0121c"
/notes="Rv0121c, (MTC1418B.03c), unknown, len: 144 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0121c"
/protein_id="CAB09449.1"
/db_xref="GI:2181963"
/db_xref="SPTREMBL:O07171"

```

misc\_feature

gene

CDS

misc\_feature

gene

CDS

```

/translation="MGEDPKLRFAQSPVARLATSTDPDTPHLVPPVVFALCARRPAEA
RDGAVYIVTAVDAKRRKTTORLRLRLANLEHNPRASVLYDSYADDWTQLMWVRADGVAIIH
RGESVRAAYRLLRAKYAQCSVPLNGPVIATVQVRWASWHA"
complement(4399..4401)
RBS
/notes="possible RBS upstream of Rv0121c"
4532..4535
RBS
/notes="possible RBS upstream of Rv0122"
4539..4907
gene
/genes="Rv0122"
4539..4907
CDS
/genes="Rv0122"
/notes="Rv0122, (MTC1418B.04), unknown, len: 133 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0122"
/protein_id="CAB09450.1"
/db_xref="GI:2181964"
/db_xref="SPTREMBL:O07172"
/translation="MAGSVSAAGIGWGLNVNTNRDQCYRVRTTVDALTHPEYRV
HTRGVQRVTRNARKHRVSKRIVAAHRHCGVPVIOEDGSLYIOGHDTSGLRLTEVVA
VEADDGLIITHAMPKEWR"
4890..4895
RBS
/genes="Rv0122"
/notes="possible RBS upstream of Rv0123"
4904..5272
gene
/genes="Rv0123"
4904..5272
CDS
/genes="Rv0123"
/notes="Rv0123, (MTC1418B.05), unknown, len: 133 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0123"
/protein_id="CAB09451.1"
/db_xref="GI:2181965"
/db_xref="SPTREMBL:O07173"
/translation="MTKKPNADYVIGDDVEVSDVLKQEEVYVDCERLTDERVEOM
ASESLAREAREANLIPGKSLSGSAHSPAVVVVSKATHAKELARSRKMSVSKL
LRPVLDFEYQRETRGILPRR"
5570..5574
RBS
/notes="possible RBS upstream of Rv0124"
5581..7044
gene
/genes="PE_PGRS"
5581..7044
CDS
/genes="PE_PGRS"
/notes="Rv0124, (MTC1418B.06), len: 487. Member of the
PGRSfamily of glycine-rich proteins, highly similar to eg.
YODP_MYCTU_Q50615 hypothetical glycine-rich 40.8 kd
protein (498 aa). fasta scores: opt: 1730 z-score: 1201.5
E(): 0, 60.7% identity in 504 aa overlap"

```

Alignment Scores:

Pred. No.:	3.52e-41	Length:	11700
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-684-215A-4 (1-132) x MTC1418B (1-11700)

Qy 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

Db 7865 ACGCCGCGTCCGATAATCTTCAGCTGTCCAGAGGAGGATTCCGCCATCCGATC 7924

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db 7925 GGGCAGGGCATGGCGATCCGGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTCAT 7984

Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60

Db 7985 ATCGGGCTACCGCTTCTCGGCTTGGGTCTTCGACACAAACGACGACGGCGCAGCA 8044

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

```

|||||
8045 GTCCAAACGCTGGTCCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGAGTG 8104
|||||
Qy 1lethrAlaValaspGlyAlaproleasnSerAlaThrAlaMetAlaAspAlaLeuasn 100
|||||
Db 8105 ATCACCGGGTCGACGGCGCTCCCATCACTCGCCACCGCATGGCGGACCGCTTAAC 8164
|||||
Qy 101 GlyHisProGlyAspValileSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 8165 GGGCATCATCCGCTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGT 8224
|||||
Qy 121 ThrGlyasnValThrLeuAlaGluGlyproProala 132
|||||
Db 8225 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 8260
|||||

RESULT 5
LOCUS AE006925 14029 bp DNA linear BCT 27-APR-2001
DEFINITION Mycobacterium tuberculosis CDC1551, section II of 280 of the
complete genome.
ACCESSION AE006925 AE000516
VERSION AE006925.1 GI:13879610
KEYWORDS
SOURCE Mycobacterium tuberculosis CDC1551
ORGANISM Mycobacterium tuberculosis CDC1551
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1 (bases 1 to 14029)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
TITLE Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 14029)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Source
1..14029
Location/Qualifiers
1..14029 "Mycobacterium tuberculosis CDC1551"
/mol_type="genomic DNA"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
complement(99..533)
/gene="MT0129"
complement(99..533)
/gene="MT0129"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44351.1"
/db_xref="GI:13879611"
/translation="MGFDPKLRFAQSPVARLATSTPDGTPHLVPVVFALGARPAEA
TGADVIYTVADAKRTTORLRRLANLEHNPASVLDSDYADDWTQLWVRADGVAAIH
RDGEVMAARLLRAKYTOQSVPLNGPVIATVQRASRHA"
complement(643..1062)
/gene="MT0130"
complement(643..1062)
/gene="MT0130"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11

product="hypothetical protein"
/protein_id="AAK44351.1"
/db_xref="GI:13879612"
/translation="MASSSLPLLRHCVSDDVTVVGVGDGDDLGCKTARRIAALVVQRA
IFLNDRTVAHSGDDAYLGHAVLPVGPVPRDPDLPHASSMYSLCMQSVNCRPPDATA
LVSVRLCHVOTDPTDSCGGRDRPQLPCAPLDYHRRH"
1047..1415
/gene="MT0131"
1047..1415
/gene="MT0131"
/note="identified by match to protein family HMM"
/transl_table=11
/product="DNA-binding protein, CopG family"
/protein_id="AAK44355.1"
/db_xref="GI:13879613"
/translation="MTKKPRNPADYVIGDVEVSDVLKOEVVVDCERLTDERVEOM
ASESLRLAREANLIPGKSLSGGSAHSPAVQVVVSKATHAKELARSKRMSVSKL
LRPVLDPEFVQRETRILPRR"
1682..3367
/gene="MT0132"
1682..3367
/gene="MT0132"
/note="similar to GB:AL123456; identified by sequence
similarity; putative"
/transl_table=11
/product="PE_PGRS family protein"
/protein_id="AAK44356.1"
/db_xref="GI:13879614"
/translation="MFAGGAGGGLRCVMFSVVAPEIVVAATDLAGIGSAISAANA
AAAAPTTAVLAAGADEVSAATAALFSGHQAAYALSAFAAFHQQFVQTLAGAGAYA
AAEAQVEQQLIAAINAPTALLRLPGINGADGAPGTQAGAGGILYNGNGSGSA
AGQAGAGGAPGLIGHGSGGAGGAGGAGGHWLWNGVGGGAGVAGVAG
GHGAGCAAGLWAGGCGGNGGADANIVSGDGLGGAGGGWLYDGGAGGHG
CGAGGAGGAGGCGGAGRGLWCTGCGAGGCGGCGTGPPLPGAGMGAAGAGG
LIGNGAGGCGGAGVAGVAGGAGNMLIGHGAGGAGGDSFANGAAGAGGAG
LHFGNGSGGHHGAVTAGNTGIGGAGGAGGADARLIGHGAGGAGGAGALVGRDGG
PGNGGAGGQLYNGGCGGCGQAGFANNIGTGTGAGNGGPAILSGNGGAG
GAGGAGGAGGCGGAGGAGGAGTGGTTLQAASVLTALFCAGPQPGDTGPG"
3519..4586
/gene="MT0133"
3519..4586
/gene="MT0133"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="serine protease, putative"
/protein_id="AAK44357.1"
/db_xref="GI:13879615"
/translation="MSNSRRSLRWSLLSVLAAVGLGLATAPAAQAPALSDORFAD
FPALPLDPSAMVAQVGPVNVNITKLGYNNAVAGTGIVIDPNGVLTNNHVIAGATD
INAFVSGGTYGVVDYDTPAVLQLRGAGLPSAATGGVAVGEPVAVANGSG
QGCTPRAPVGRVVALQCTVOASDSLTAETLNGLIQFDAAIQPGSGGVPVNLGQ
VWMNTAASDNFQSLGQGGQFAIPIGQAMATAGQIRSGGSGPTVHIGPTAFGLGVVD
NNGCARVQRVVGSAAPASLGISTGDTTAVDGPAPINSATAMADALNHHHPGDIISVT
WQTKSGGTRTGNVTLAEGPPA"
4695..6500
/gene="MT0134"
4695..6500
/gene="MT0134"
/note="similar to GP:2808807; identified by sequence
similarity; putative"
/transl_table=11
/product="alpha-amylase family protein"
/protein_id="AAK44358.1"
/db_xref="GI:13879616"
/translation="MNEAHSVEHPVQGSHEVGVVHEDPAKDFGSAALPADPTWF
KHAFYEVLYRAFFDASADGDLRLDLDYQLWLGLDLCIWLPPFVDSPLRDGGVD
IRDFYKVLPEFGTVDDDFVALDAHRGIRITLDVMNHTSESHPWQESRRDPGPI
GDYVWSDTSEYTDARIIFVDTEESNWSDEPVRQRQFTWHRFFSHQPDLDNPAVOE
AMIDVIRFWLGLDGRFLDAVPYLFEREECTNENLPETHAFLKRVKRVVDDEPGRV

```

LLAENANQPDVVEYFGDPNTGGDECHMAFHFFLMPRIFMVAVRRESRFPISSEIAOTI  
 PIQMAQGFIFLRNDELTLMEVYTDERDYMAYEAKDPRMKANVGIRRRALPLLDN  
 RQETETALLSLSPGSPVLYYDEIGMGDVIWLGDRDGVRIQWOTPPDRNAGFSTAN  
 PGRLYLPSSQDPVYQAVNVEAORDTSTLLNTRMLAVRRRHFAFAVGAFOELGG  
 SNPSVLAVROVAGDGDGTIVLNVNLSRFPOPLELDLQWNTYTPVELTGHVFEPRIG  
 QVPYLLTLPGRHGFYWFQTLTHEVGAPTCGGERRL"  
 6603..7970  
 /gene..WT0135"  
 /gene..WT0135"  
 6603..7970  
 /gene..WT0135"  
 /note~similar to GP:2808801; identified by sequence  
 similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="pep2 protein"  
 /protein\_id="AAK44359.1"  
 /db\_xref="GI:13879617"  
 /translation="MTRSDTLATKPLMSDMLPRQWAGNRELATVKGVVVALRH  
 LDVLVDVYTDGATERGYQVLVGHDFPASEYGTKKAIGVADDTGFDALYDVAGQF  
 LLSLVSAVGCTGTEVTFREDPLFPAAQPRVCDAESNTSVIFDRRAILKQFR  
 RVSGINPDIENLVLRAGNPHVRLLAGAYQGRPNRSPTDALATLVGMYTEYANA  
 AEGWAMATASVDLYFAEGDLYAHEVGDFAGESYRLGEAVASVHATLADSLGTAQATF  
 PDMRLARKLVSTAVVPELRYAETIQEQFKLAAEAATVORVHGDLHLGOVLRTPE  
 WLLIDFECEQDIPERRADPSLPURDVAGVLRSEFAAYGVLVDQATDKQLAARARE  
 VERNRAFCDSYAVASGIDPRDSALLGAYELDKAVYETGYETRRHPGLPIPLRSIA  
 RLTAAS"  
 8038..8817  
 /gene..WT0136"  
 /gene..WT0136"  
 8038..8817  
 /gene..WT0136"  
 /note~identified by Glimmer2; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAK44360.1"  
 /db\_xref="GI:13879618"  
 /translation="MQREIYDGEARLSWLAALAGILGATAFTHSAGYFVTMTGNSQ  
 RAVLGLGDDAWMSVTASLLILDFVAGVVTASVCRHFHFAAHPGTVTLTFSLLFA  
 GVDIMLGWHRESMLDFVPIILFVFGIGALNTSFVKDGEVSPLYVTGLVKNMGQIE  
 RHLGKGVDEWLGTFLLHASFVIGAAAGGALISMVYTGQMLVAVAAVCAATTGTYTLH  
 ADRLGYNQKRQPGKRLFRALRGELDSGTSTPATNYGSS"  
 complement(8949..9989)  
 /gene..WT0137"  
 complement(8949..9989)  
 /gene..WT0137"  
 /note~similar to SP:P31953 GB:X57229 PID:48828  
 PID:1567734 PID:1877254; identified by sequence  
 similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="esterase, putative, antigen 85-C"  
 /protein\_id="AAK44361.1"  
 /db\_xref="GI:13879619"  
 /translation="MGLVAVMTFPEQVRRLRSATTLPRRLAIAMGAVILVGLVGTFF  
 GGPATAGAFSRPLVPEYLQVPSASMGRIKVFQGGPHAVYLLDLGLRAQDDYNGMD  
 INTAFPEYVYQSGLSVIMPYGGSSFTDYQYQSQSNQNTYKWTETFLTREMPALQ  
 ANKVYPTGNAAGVLSMSGSSALLTAAAYQOPPYAASLSGFLNPGSGWPTLIGLAM  
 NDSGYNANSMWGPSDDPAWKRNDPMVQIPLRVANTRIMWYCGNTPSPDLGDNTPA  
 KFLGLTLRTNOTFRDYIYAADGGNGVFNPPNGTSHWPYNQQLVAMKRAIDQHVLN  
 ATPAPAAPAA"  
 10218..10673  
 /gene..WT0138"  
 10218..10673  
 /gene..WT0138"  
 /note~similar to SP:P08634 PID:46222; identified by  
 sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="modulation protein N-related protein"  
 /protein\_id="AAK44362.1"  
 /db\_xref="GI:13879620"

Alignment Scores:

Pred. No.:	4.21e-41	Length:	14029
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0
US-09-684-215A-4 (1-132) * AE006925 (1-14029)			
Qy	1	ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile	20
Db	4188	ACGGCCCGCTGCCATTAACCTCCAGCTGTCGCCAGGTGGCAGGGATTCGCCATTCGCATC	4247
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHAs	40
Db	4248	GGGCAGCGCATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTACCCACCGTTCAT	4307
Qy	41	IleGlyProThrAlaPheLeuGlyLeuGlyValValaIaSPAsnGlnGlyAlaAArg	60
Db	4308	ATCGGGCTTACGCCCTTCCTGGCTGGGTGTTCGACAAACGCAACGGCAGCGCACGA	4367
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	4368	GTCCAACGCGCTGCTCGGGAGCGCTCCGGCGGCAAGTCTCGGCAATCCACCGGCGACGTG	4427
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaSPAlaLeuAsn	100
Db	4428	ATCACCGCGGTGCACCGCGCTCCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAC	4487
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg	120
Db	4488	GGCGATCATCCGGTGCACGTCACTCGGTGACCTGCCAAACCAAGTCGGGGCGCAGCGT	4547
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132
Db	4548	ACAGGAAACGTGCATTTGCCGAGGACGCCCGGCC	4583
RESULT 6			
LOCUS	BX248334	343050 bp	DNA linear BCT 11-JUN-2003
DEFINITION	Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment 1/14.		
ACCESSION	BX248334		
VERSION	BX248334.1		
KEYWORDS	complete genome.		
SOURCE	Mycobacterium bovis subsp. bovis AF2122/97		
ORGANISM	Mycobacterium bovis subsp. bovis AF2122/97		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterinae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
AUTHORS	Garnier,T., Elgmeler,K., Camus,J.-C., Medina,N., Mansoor,H., Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsemp,C., Simon,S., Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L., Wheeler,P.R., Parkhill,J., Barrell,H.G., Cole,S.T., Gordon,S.V. and Hewinson,G.		
TITLE	The complete genome sequence of Mycobacterium bovis		
JOURNAL	Online Publication		
REMARK	PNAS 10.1073/pnas.1130426100 ( Microbiology )		
REFERENCE	2 (bases 1 to 343050)		
AUTHORS	Garnier,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex 15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the Mycobacterium bovis sequencing teams, TB Research Group, Veterinary Laboratories Agency Weybridge, Woodham Lane, New Haw, Addlestone, Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus, Hinxton,Cambridge CB10 1SA, UK. Pr4 Annotation, Genopole, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France. Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France		
FEATURES	Location/Qualifiers		

## FEATURES

```

source
1. .343050
/organism="Mycobacterium bovis subsp. bovis AF2122/97"
/mol_type="genomic DNA"
/strain="AF2122/97"
/db_xref="taxon:233413"
1. .1524
/gene="dnaA"
/locus_tag="Mb0001"
/locus_tag="Mb0001"
/note="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001, len: 507 aa, from Mycobacterium tuberculosis strain H37Rv, (99.6% identity in 507 aa overlap). dnaA, chromosomal replication initiator protein (see citations below), equivalent to other Mycobacterial CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g. P46388|DnaA_MYCLE from Mycobacterium leprae (502 aa); Q9L7L7|DnaA_MYCPA from Mycobacterium paratuberculosis (509 aa); P49990|DnaA_MYCAV from Mycobacterium avium (508 aa); P49992|DnaA_MYCSM from Mycobacterium smegmatis (504 aa); etc. Also highly similar to others except in N-terminus e.g. Q9ZH75|DnaA_STRCH CHROMOSOMAL REPLICATION INITIATOR PROTEIN from Streptomyces chrysomallus (624 aa); Q9ZH76|DnaA_STRRE from Streptomyces reticuli (643 aa); DnaA_ECOLI|P03004|B3702 chromosomal replication initiator protein from Escherichia coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0, (43.2% identity in 389 aa overlap); etc. Contains PS00017 ATP/GTP-binding site motif A (P-loop) and PS01008 DnaA protein signature. BELONGS TO THE DnaA FAMILY. Note that the first base of this gene has been taken as base 1 of the Mycobacterium bovis genomic sequence."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DnaA"
/protein_id="CAD92863.1"
/db_xref="GI:31616763"
/translation="MTDDPGSGFTTVNNAVSVSELNGDPKVDGPPSDANLSAPLTPOQ RAWLNVLQPLTIVEGFALLSPSSFQVNERHLRAPITDALSRRLGHQIQLVRIAP PATDEADTTVPSENPATSPDTTNDDEIDSAAGAAGNQHSWFSPTFRPNRTDS ATAGVTSNNRRYTFDFVGASRNFAHAALAATAEPARAINPFLFIWGESGLGRKTHLL HAAGNYAQRLFGMRKYVYSTEEDFNISLRDRKRVAKRSYRDVVLLDVDDIQPI EGKEGTQEFFFFHTFYHNANKOIVISSDRPPQLATLEDRLTRFEWGLTVDQPPE LETAIRAIKKQAQMELRAIPDDVLFHIAASSIERNIIELEGALIRVTAFASINKTPIDK ALAEIVRLDIADANPMQISAATIMAAATEYFTTVEELRGPKTKALACSOIAMYL CRELDLDSLKPKGFGQRDHFTVMYAQRKILSEMAERREVFHDVKELTTIRQRSKR" 2052..3260
/gene="dnaN"
/locus_tag="Mb0002"
/locus_tag="Mb0002"
2052..3260
/gene="dnaN"
/locus_tag="Mb0002"
/BC_number="2.7.7"
/note="Mb0002, dnaN, len: 402 aa. Equivalent to Rv0002, len: 402 aa, from Mycobacterium tuberculosis strain H37Rv, (99.8% identity in 402 aa overlap). dnaN, DNA polymerase III (beta chain) (EC 2.7.7) (see citations below), equivalent to other Mycobacterial DNA POLYMERASES III BETA CHAIN e.g. NP_301130.1|NC_002677 from Mycobacterium leprae (399 aa); Q9L7L6|DP3B_MYCPA from Mycobacterium avium subsp. paratuberculosis (399 aa); P52851|DP3B_MYCSM from Mycobacterium smegmatis (397 aa); etc. Also highly similar to others e.g. P27903|DP3B_STRCO DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus luteus (310 aa); P52023|DP3B_SYNPF from Synecococcus sp. strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in neighbouring cosmid MTCY10H4.01."
/codon_start=1
/evidence=experimental
/transl_table=11

```



```
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AR182442 (1-447)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGCCGGCTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyClnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCCGATCGGGTGGGGTCAACCCCGTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCTACCGCTTCCTCGGCTGGGTGTTCGACAAACAAACGGCAACGGCGCACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGTCCGATCGCGGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGCTTAAC 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120
Db 311 GGGCATCATCCGCTTCCTCGGCTGGGTGTTCGACAAACAAACGGCAACGGCGCACGTG 370
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGAACGTGACATTTGGCCGAGGACCCCGGCC 406
RESULT 9
LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6350456.
ACCESSION AR194825
VERSION AR194825.1 GI:20244262
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.24e-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AR194825 (1-447)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGCCGGCTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyClnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCCGATCGGGTGGGGTCAACCCCGTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCTACCGCTTCCTCGGCTGGGTGTTCGACAAACAAACGGCAACGGCGCACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGTCCGATCGCGGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGCTTAAC 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120
Db 311 GGGCATCATCCGCTTCCTCGGCTGGGTGTTCGACAAACAAACGGCAACGGCGCACGTG 370
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGAACGTGACATTTGGCCGAGGACCCCGGCC 406
RESULT 9
LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6350456.
ACCESSION AR194825
VERSION AR194825.1 GI:20244262
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.24e-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AR233097 (1-447)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGCCGGCTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyClnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCCGATCGGGTGGGGTCAACCCCGTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCTACCGCTTCCTCGGCTGGGTGTTCGACAAACAAACGGCAACGGCGCACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGTCCGATCGCGGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGCTTAAC 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120
Db 311 GGGCATCATCCGCTTCCTCGGCTGGGTGTTCGACAAACAAACGGCAACGGCGCACGTG 370
```

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|||||  
Db 371 ACAGGGAAGTGCATTTGGCCGAGGACCCCGGCC 406

## RESULT 11

AX429596  
LOCUS AX429596 447 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 4 from Patent EP1203817.  
ACCESSION AX429596  
VERSION AX429596.1 GI:21540845

KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.

## REFERENCE

1 Reed, S.G., Skeiky, Y.A., Dillon, D.C., Campos-Neto, A., Houghton, R.L.,  
Vedvick, T.S., and Twardzik, D.R.  
Compounds and methods for immunotherapy and diagnosis of  
tuberculosis

JOURNAL Patent: EP 1203817-A 4 08-MAY-2002;  
CORIXA CORPORATION (US)

FEATURES  
source  
1. 447  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3.24e-42 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x AX429596 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACAGGCGGTCGATTAACCTCCAGCTGTCCAGGTTGGCGGAGGATTCGCCATTCGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTCAT 130  
QY 41 IleGlyProThrAlaPheLeuGlyValValAlaSerLeuGlyIleSerThrGlyAspVal 60  
|||||  
Db 131 ATCGGCGCTACCGCTTCCGCGTGGGTGTGTCGACAAACGCGGCGGCGACGA 190  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||  
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGGGGCAAGTCTCGGATCTCCACCGCGGACGTG 250  
QY 81 IleThrAlaValAlaGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|||||  
Db 251 ATCAGCGGCGTGGAGCGGCTCCGATCAACTCGGCGGCGGCGGCGGCGGCGGCGG 310  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120  
|||||  
Db 311 GGGCATCATCCCGGTGACGTATCTCGGTGAACCTCGGCGGCGGCGGCGGCGGCGG 370  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|||||  
Db 371 ACAGGGAAGTGCATTTGGCCGAGGACCCCGGCC 406

## RESULT 12

BD006325  
LOCUS BD006325 447 bp DNA linear PAT 31-JAN-2002  
DEFINITION Compounds and methods for diagnosis of Tuberculosis.  
ACCESSION BD006325  
VERSION BD006325.1 GI:18634696

## KEYWORDS

SOURCE JP 2001500383-A/4.  
ORGANISM unidentified;  
unclassified;

## REFERENCE

1 (bases 1 to 447)

## AUTHORS

Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,  
Vedvick, T.S., Twardzik, D.R. and Lodes, M.J.

## TITLE

Compounds and methods for diagnosis of Tuberculosis

## JOURNAL

Patent: JP 2001500383-A 4 16-JAN-2001;

## COMMENT

CORIXA CORP

OS Unidentified

PN JP 2001500383-A/4

PD 16-JAN-2001

PF 07-OCT-1997 JP 1998518432

PR 11-OCT-1996 US 08/729622.13-MAR-1997 US 08/818111 PI

STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS

NETO, RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R THWARDZIK, PI

MICHAEL J LODES

PC C12N15/31, C07K14/35, C07K16/12, C12Q1/68, C12N15/62, G01N33/53 CC

Strandedness: Single;

CC Topology: Linear;

FH Key Location/Qualifiers

FT source 1. 447

/organism="unidentified".

Location/Qualifiers

1. 447

/organism="unidentified"

/mol\_type="genomic DNA"

/db\_xref="taxon:32644"

BASE COUNT 79 a 146 c 149 g 72 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 3.24e-42 Length: 447

Score: 670.00 Matches: 131

Percent Similarity: 99.24% Conservative: 0

Best Local Similarity: 99.24% Mismatches: 1

Query Match: 99.26% Indels: 0

DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x BD006325 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

|||||

Db 11 ACAGGCGGTCGATTAACCTCCAGCTGTCCAGGTTGGCGGAGGATTCGCCATTCGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

|||||

Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTCAT 130

QY 41 IleGlyProThrAlaPheLeuGlyValValAlaSerLeuGlyIleSerThrGlyAspVal 60

|||||

Db 131 ATCGGCGCTACCGCTTCCGCGTGGGTGTGTCGACAAACGCGGCGGCGACGA 190

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

|||||

Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGGGGCAAGTCTCGGATCTCCACCGCGGACGTG 250

QY 81 IleThrAlaValAlaGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

|||||

Db 251 ATCAGCGGCGTGGAGCGGCTCCGATCAACTCGGCGGCGGCGGCGGCGGCGGCGG 310

QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120

|||||

Db 311 GGGCATCATCCCGGTGACGTATCTCGGTGAACCTCGGCGGCGGCGGCGGCGGCGG 370

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

|||||

Db 371 ACAGGGAAGTGCATTTGGCCGAGGACCCCGGCC 406

RESULT 13

BD006445



```

LOCUS      BD006445                      447 bp    DNA    linear    PAT 31-JAN-2002
DEFINITION Compounds and methods for immunotherapy and diagnosis of
Tuberculosis.
ACCESSION  BD006445
VERSION    BD006445.1 GI:18634816
KEYWORDS   JP 2001501832-A/4.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 447)
AUTHORS    Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
TITLE      Compounds and methods for immunotherapy and diagnosis of
JOURNAL    Patent: JP 2001501832-A 4 13-FEB-2001;
CORIXA CORP
COMMENT     OS Unidentified
            PN JP 2001501832-A/4
            PD 13-FEB-2001
            PF 07-OCT-1997 JP 1998518456
            PR 11-OCT-1996 US 08/730510,13-MAR-1997 US 08/818112 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI ANTONIO CAMPOS
            NETO,
            PI RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC C12N15/31,C07K14/35,A61K39/04,A61K48/00,A61K49/00,C12N15/62,
            C07K19/00,
            PC G01N33/50,G01N33/60,G01N33/569,C12N1/19,C12N1/20,C12N1/21, PC
            C12N5/10//
            PC (C12N1/21,C12R1:19)
            CC Strandedness: Single;
            CC Topology: Linear;
            FH Key Location/Qualifiers
            FT source 1..447
            /organism="Unidentified".
FEATURES             Location/Qualifiers
     . source        1..447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
BASE COUNT          79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.:          3,24e-42      Length:      447
Score:              670.00      Matches:    131
Percent Similarity: 99.24%      Conservative: 0
Best Local Similarity: 99.24%   Mismatches: 1
Query Match:        99.26%      Indels:    0
DB:                 6          Gaps:      0

US-09-684-215A-4 (1-132) x BD006445 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGCCAGGGTGGGCGAGGATTCGCATTCGCATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 71 GGGCAGCGGATGCGATCGCGGCCAGATCCGATCGGGTGGGGTCCACCCCGTTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCACAAACGCGACGCGCACGA 130
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCACGCGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGAGGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGGTGACGCGCTCCGATCACTCCGCCACCGCATGCGGACGCGCTTAC 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

```

```

Db 311 GGCATCATCCCGTACGTCATCTCGTGAACCTGGCAACCAAGTCGGCGGCACCGCT 370
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGGAACGTGACATTTGCCGAGGACCCCGGCC 406

RESULT 14
BD069285
LOCUS      BD069285                      447 bp    DNA    linear    PAT 27-AUG-2002
DEFINITION Compounds and methods for immunotherapy and diagnosis of
Tuberculosis.
ACCESSION  BD069285
VERSION    BD069285.1 GI:22614888
KEYWORDS   JP 2001517069-A/4.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 447)
AUTHORS    Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H. and Twardzik,D.R.
TITLE      Compounds and methods for immunotherapy and diagnosis of
JOURNAL    Patent: JP 2001517069-A 4 02-OCT-2001;
CORIXA CORP
COMMENT     OS Unidentified
            PN JP 2001517069-A/4
            PD 02-OCT-2001
            PF 30-AUG-1996 JP 1997511464
            PR 01-SEP-1995 US 08/523436, 22-SEP-1995 US 08/533634 PR
            22-MAR-1996 US 08/620874,05-JUN-1996 US 08/659683 PR
            12-JUL-1996 US 08/680574
            PI STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
            PI NETO,
            PI RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
            C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
            C12N5/10,
            PC C12N1/21//A61K39/04,(C12N1/21,C12R1:19)
            CC Strandedness: Single;
            CC Topology: Linear;
            CC Compounds and methods for immunotherapy and diagnosis of CC
            tuberculosis
FEATURES             Location/Qualifiers
     FH Key          Location/Qualifiers
     FT source       1..447
                     /organism="Unidentified".
     . source        1..447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
BASE COUNT          79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.:          3,24e-42      Length:      447
Score:              670.00      Matches:    131
Percent Similarity: 99.24%      Conservative: 0
Best Local Similarity: 99.24%   Mismatches: 1
Query Match:        99.26%      Indels:    0
DB:                 6          Gaps:      0

US-09-684-215A-4 (1-132) x BD069285 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGCCAGGGTGGGCGAGGATTCGCATTCGCATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 71 GGGCAGCGGATGCGATCGCGGCCAGATCCGATCGGGTGGGGTCCACCCCGTTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCACAAACGCGACGCGCACGA 190

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:28:59 ; Search time 70.5517 Seconds  
(without alignments)  
825.814 Million cell updates/sec

Title: us-09-684-215a-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGGGGAIFI.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlx  
-O=/cgn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_593/app\_query.fasta\_1.853  
-DB-Issued\_Patents\_NA -QWTF=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -CGN\_1\_1\_145 @runat\_05092003\_072203\_593 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	675	100.0	2287	US-09-223-040-1	Sequence 1, Appli
2	675	100.0	4403765	US-09-103-840A-2	Sequence 2, Appli
3	675	100.0	4411529	US-09-103-840A-1	Sequence 1, Appli
4	670	99.3	447	US-08-818-112-4	Sequence 4, Appli
5	670	99.3	447	US-08-818-111-4	Sequence 4, Appli
6	670	99.3	447	US-09-056-556-4	Sequence 4, Appli
7	670	99.3	447	US-09-072-596-4	Sequence 4, Appli
8	670	99.3	1872	US-08-818-112-17	Sequence 17, Appl
9	670	99.3	1872	US-08-818-111-17	Sequence 17, Appl
10	670	99.3	1872	US-09-056-556-17	Sequence 17, Appl
11	670	99.3	1872	US-09-072-596-17	Sequence 17, Appl
12	632	93.6	822	US-09-736-457-1862	Sequence 1862, Ap

13	632	93.6	900	4	US-09-643-597-353	Sequence 353, App
14	632	93.6	900	4	US-09-606-421B-353	Sequence 353, App
15	632	93.6	945	4	US-09-736-457-1861	Sequence 1861, Ap
16	632	93.6	1012	4	US-09-643-597-351	Sequence 351, App
17	632	93.6	1012	4	US-09-606-421B-351	Sequence 351, App
18	632	93.6	1464	4	US-09-620-412C-348	Sequence 348, App
19	632	93.6	1464	4	US-09-598-419-348	Sequence 348, App
20	632	93.6	1557	4	US-09-620-412C-332	Sequence 332, App
21	632	93.6	1557	4	US-09-598-419-332	Sequence 332, App
22	632	93.6	1578	4	US-09-556-877-188	Sequence 188, App
23	632	93.6	1578	4	US-09-620-412C-188	Sequence 188, App
24	632	93.6	1578	4	US-09-598-419-188	Sequence 188, App
25	632	93.6	1752	4	US-09-620-412C-352	Sequence 352, App
26	632	93.6	1752	4	US-09-598-419-352	Sequence 352, App
27	632	93.6	1758	4	US-09-620-412C-336	Sequence 336, App
28	632	93.6	1758	4	US-09-598-419-336	Sequence 336, App
29	632	93.6	1860	4	US-09-620-412C-308	Sequence 308, App
30	632	93.6	1860	4	US-09-598-419-308	Sequence 308, App
31	632	93.6	1896	4	US-09-620-412C-324	Sequence 324, App
32	632	93.6	1896	4	US-09-598-419-324	Sequence 324, App
33	632	93.6	1941	4	US-09-620-412C-316	Sequence 316, App
34	632	93.6	1941	4	US-09-598-419-316	Sequence 316, App
35	632	93.6	1965	4	US-09-620-412C-340	Sequence 340, App
36	632	93.6	1965	4	US-09-598-419-340	Sequence 340, App
37	632	93.6	2052	4	US-09-620-412C-356	Sequence 356, App
38	632	93.6	2052	4	US-09-598-419-356	Sequence 356, App
39	632	93.6	2076	4	US-09-620-412C-312	Sequence 312, App
40	632	93.6	2076	4	US-09-598-419-312	Sequence 312, App
41	632	93.6	2103	4	US-09-620-412C-344	Sequence 344, App
42	632	93.6	2103	4	US-09-598-419-344	Sequence 344, App
43	632	93.6	2148	4	US-09-620-412C-320	Sequence 320, App
44	632	93.6	2148	4	US-09-598-419-320	Sequence 320, App
45	632	93.6	2148	4	US-09-620-412C-328	Sequence 328, App
					US-09-598-419-320	Sequence 320, App

#### ALIGNMENTS

#### RESULT 1

; Sequence 1, Application US/09223040  
; Patent No. 6544522  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014038-0090100S  
; CURRENT APPLICATION NUMBER: US/09/223,040  
; CURRENT FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2287  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion  
; OTHER INFORMATION: protein Ra12-TbH9-Ra35  
; NAME/KEY: modified\_base  
; LOCATION: (30)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: modified\_base  
; LOCATION: (33)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: CDS  
; LOCATION: (42)..(2231)  
; NAME/KEY: modified\_base  
; LOCATION: (2270)  
; OTHER INFORMATION: n = g, a, c or t  
US-09-223-040-1

## Alignment Scores:

Pred. No.: 1.85e-61 Length: 2287  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-223-040-1 (1-2287)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 63 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 122  
|||||  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 123 GGGCAGGCGATGGCGATCGCGGCCAGATCCCATCGGTGGGGTCCACCCACCGTTTAT 182  
|||||  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|||||  
Db 183 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGTCACAAACGCAACGCGCACGA 242  
|||||  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||  
Db 243 GTCCACAGCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 302  
|||||  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|||||  
Db 303 ATCAGCGGCTCGACGGCGCTCCGATCACTCGGCCACCGCGATGGCGACCGCTTAAC 362  
|||||  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|||||  
Db 363 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCT 422  
|||||  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|||||  
Db 423 ACAGGAACTGACATTTGGCCAGGAGGACCCCGGCC 458  
|||||

## RESULT 2

US-09-103-840A-2

; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

## Alignment Scores:

Pred. No.: 2.7e-57 Length: 4403765  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-103-840A-2 (1-4403765)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 151984 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 152043  
|||||  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 152044 GGGCAGGCGATGGCGATCGCGGCCAGATCCCATCGGTGGGGTCCACCCACCGTTTAT 152103  
|||||  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|||||  
Db 152104 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGTCGACAAACGCAACGCGCACGA 152163  
|||||  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||  
Db 152164 GTCCACAGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 152223  
|||||  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|||||  
Db 152224 ATCAGCGGCTCGACGGCGCTCCGATCACTCGGCCACCGCATGGCGACCGCTTAAC 152283  
|||||  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|||||  
Db 152284 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCT 152343  
|||||  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|||||  
Db 152344 ACAGGAACTGACATTTGGCCAGGAGGACCCCGGCC 152379  
|||||

## RESULT 3

US-09-103-840A-1

; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

## Alignment Scores:

Pred. No.: 2.7e-57 Length: 4411529  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-103-840A-1 (1-4411529)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 151815 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 151874  
|||||  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 151875 GGGCAGGCGATGGCGATCGCGGCCAGATCCCATCGGTGGGGTCCACCCACCGTTTAT 151934  
|||||  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|||||  
Db 151935 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGTCGACAAACGCAACGCGCACGA 151994  
|||||



```
Alignment Scores:
Pred. No.: 7.8e-62 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-08-818-111-4 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTGCCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 71 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCAT 130
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB 131 ATCGGCCCTACCGCTTCTCGGCTTGGGTCTGTGCGACAAACACGCGCACGCGCACGA 190
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 191 GTCCAAACCGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 251 ATCACC CGGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 310
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 311 GGGCATCATCCGCTGAGTCTCATCTCGGTGAACCTGGCAACCAAGTCGGCGGCGACGGT 370
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 371 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 406

RESULT 6
US-09-056-556-4
; Sequence 4, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; TREATMENT OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-056-556-4
Alignment Scores:
Pred. No.: 7.8e-62 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-056-556-4 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTGCCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 71 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCAT 130
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB 131 ATCGGCCCTACCGCTTCTCGGCTTGGGTCTGTGCGACAAACACGCGCACGCGCACGA 190
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 191 GTCCAAACCGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 251 ATCACC CGGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 310
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 311 GGGCATCATCCGCTGAGTCTCATCTCGGTGAACCTGGCAACCAAGTCGGCGGCGACGGT 370
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 371 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 406

RESULT 7
US-09-072-596-4
; Sequence 4, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-596-4

Alignment Scores:  
Pred. No.: 7.8e-62 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-072-596-4 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGCCCGCGTCCGATACTTCCAGCTGCCAGGGTGGGCGAGGATTCGCCATTCGCATC 70  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 71 GGCAGCGCATGGCGATCGCGGCCACATCCGATCGGGTGGGGGTCAACCACCGTTCAT 130  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
Db 131 ATCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACACGCGGCGCACGA 190  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCACCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGGCGGTG 250  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGCGTTCGACGCGCTCCGATCACTCGGCCACCGGATGGCGGCGCTTAAC 310  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 311 GGCATCATCCGGTGACGTCTCGGTGAACCTGGCAACCAAGTCGGGCGGCGACGCGT 370  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 371 ACAGGAAACGTGACATTTGGCGGAGGACCCCGGCC 406

## RESULT 8

US-08-818-112-17

Sequence 17, Application US/08818112

Patent No. 6290969

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

APPLICANT: Campos-Neto, Antonio

APPLICANT: Houghton, Raymond

APPLICANT: Twardzik, Daniel R.

APPLICANT: Twardzik, Thomas S.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7052  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/818,112  
APPLICATION NUMBER: US/08/818,112  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-112-17

Alignment Scores:  
Pred. No.: 4.79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-4 (1-132) x US-08-818-112-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGCCCGCGTCCGATACTTCCAGCTGCCAGGGTGGGCGAGGATTCGCCATTCGCATC 817  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 818 GGCAGCGCATGGCGATCGCGGCCAAATCCGATCGGGTGGGGGTCAACCACCGTTCAT 877  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
Db 878 ATCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACACGCGGCGCACGA 937  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCACCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGCGAGGTG 997  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCACCGCGTTCGACGCGCTCCGATCACTCGGCCACCGGATGGCGGCGCTTAAC 1057  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 1058 GGCATCATCCGGTGACGTCTCGGTGAACCTGGCAACCAAGTCGGGCGGCGACGCGT 1117  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 1118 ACAGGAAACGTGACATTTGGCGGAGGACCCCGGCC 1153

## RESULT 9

US-08-818-111-17

Sequence 17, Application US/08818111

Patent No. 6338852

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,111  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-111-17

Alignment Scores:  
Pred. No.: 4,79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-08-818-111-17 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 758 ACGGGCGCGTCCGATTAACCTCCAGGTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGATGGCGATCCGGGCCAATCCATCGGTGGGGGTCCACCCACCGTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 878 ATCGGGCTACCGCTCTCCGCTGGGTGGTGTTCGACAAACACGCGACGCACCA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCACACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGCTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 998 ATCAGCGCGTGCAGCGCTCCGATCAACTCGGCCACCGCATCTCCACCGCGACGCTTAC 1057  
QY 101 GlyHisAsProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 1058 GGGCATCATCCGGTGCAGTCACTCTCGGTGACTGGCAACCAAGTCCGGCGCGACGCT 1117

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 1118 ACAGGGAAGTGCATTTGGCCGAGGAGCCCGCGCC 1153

RESULT 10  
US-09-056-556-17  
Sequence 17, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-056-556-17

Alignment Scores:  
Pred. No.: 4,79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-056-556-17 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 758 ACGGGCGCGTCCGATTAACCTCCAGGTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGATGGCGATCCGGGCCAATCCATCGGTGGGGGTCCACCCACCGTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 878 ATCGGGCTACCGCTCTCCGCTGGGTGGTGTTCGACAAACACGCGACGCACCA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCACACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGCTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100



Db 998 ATCAGCGGGTCGACGGCGCTCCGATCAATCGGCCACCAGGATGGCGGACGCGCTTAAC 1057  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|||||  
Db 1058 GGGCATCATCCGGTGACGTATCTCGGTGAACCTGGCAACCAAGTCGGGGGGACGCGT 1117  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProLa 132  
|||||  
Db 1118 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 1153  
RESULT 11  
US-09-072-596-17  
; Sequence 17, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Twardzik, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESSES:  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Makl, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-072-596-17  
Alignment Scores:  
Pred. No.: 4.79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0  
US-09-684-215A-4 (1-132) x US-09-072-596-17 (1-1872)  
Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 758 ACGGCCGCTCCGATAAATTCAGCTGTCCAGGGTGGCAGGAGGATTCGCCATTCGATC 817  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
|||||

Db 818 GGGCAGCGATGCGGATCGCGGGCAAATCCGATCGGGTGGGGGTACCCACCGCTTCAT 877  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
|||||  
Db 878 ATCGGGCTACCGCTTCTCTCGGTGGTGTTCGACAAACACGGCAACGGCGCACGA 937  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||  
Db 938 GTCCAACGCGTGGTCGGAAGCGCTCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 997  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|||||  
Db 998 ATCAGCGGGTCGACGGCGCTCCGATCAATCGGCCACCAGGATGGCGGACGCGCTTAAC 1057  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|||||  
Db 1058 GGGCATCATCCGGTGACGTATCTCGGTGAACCTGGCAACCAAGTCGGGGGGACGCGT 1117  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProLa 132  
|||||  
Db 1118 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 1153  
RESULT 12  
US-09-736-457-1862  
; Sequence 1862, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fangner, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1862  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1862  
Alignment Scores:  
Pred. No.: 1.59e-57 Length: 822  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 4 Gaps: 1  
US-09-684-215A-4 (1-132) x US-09-736-457-1862 (1-822)  
Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCCGCTCCGATAAATTCAGCTGTCCAGGGTGGCAGGAGGATTCGCCATTCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
|||||  
Db 82 GGGCAGGCGATGGCGATCGCGGGCAGATCAAG-----CTTCCACCGCTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
|||||  
Db 130 ATCGGGCTACCGCTTCTCTCGGTGGTGTTCGACAAACACGGCAACGGCGCACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||

Db 190 GTCCAAACGGCTCGGGAGCGCTCCGGCGGAAGTCTCGGATCTCCACCGCGACGCTG 249  
QY 81 ILeThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCAGCGCGCTCGAGCGCTCCGATCAACTCGGCACCGCATGGCGGCGGCTTAAAC 309  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCT 369  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405

## RESULT 13

US-09-643-597-353  
; Sequence 353, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-643-597-353

Alignment Scores:  
Pred. No.: 1.78e-57 Length: 900  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 4 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-643-597-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGCTCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAG-----CTTCCCGCTTCAT 129  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 130 ATCGGCGCTTACCGCTTCTCCGCTTGGGTGTGTGTCGACAAACGCGGCGACGCA 189  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCAAACGGCTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGGCGACG 249  
QY 81 ILeThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCAGCGCGCTCGAGCGCTCCGATCAACTCGGCACCGCATGGCGGCGGCTTAAAC 309

QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCT 369  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405

## RESULT 14

US-09-606-421B-353  
; Sequence 353, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-606-421B-353

Alignment Scores:  
Pred. No.: 1.78e-57 Length: 900  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 4 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-606-421B-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGCTCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAG-----CTTCCCGCTTCAT 129  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 130 ATCGGCGCTTACCGCTTCTCCGCTTGGGTGTGTGTCGACAAACGCGGCGACGCA 189  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCAAACGGCTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGGCGACG 249  
QY 81 ILeThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCAGCGCGCTCGAGCGCTCCGATCAACTCGGCACCGCATGGCGGCGGCTTAAAC 309  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCT 369  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405

## RESULT 15

US-09-736-457-1861  
; Sequence 1861, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736.457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1861  
; LENGTH: 945  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1861

Alignment Scores:  
Pred. No.: 1.9e-57 Length: 945  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
Db: 4 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-736-457-1861 (1-945)

Qy	1	ThrAlaAspAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	22	ACGGCCGGTCCGATACTTCAGCTGCCAGGTGGCGAGGATTCGCCATCCCGATC	81
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40
Db	82	GGCGAGGCGATGGCGATCGCGGCCAGATCAAG-----CTTCCACCGTTTCAT	129
Qy	41	IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAlaArg	60
Db	130	ATCGGGCCCTACCGCTTCCTCGCTTGGTGTGTTCGACAAACACGGCAGCGCACGA	189
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	190	GTCCAACCGGTGGTTCGGAGGCGTCCGGCGGCAAGTCTCGCATCTCCACGGCGACGTG	249
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	250	ATCACCGGGTTCGACGGGGTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAAC	309
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
Db	310	GGGCATCATCCGGTGGAGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGCACGCGT	369
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132
Db	370	ACAGGGAACGTGACATTGGCCGAGGGACCCCGGGCC	405

Search completed: September 5, 2003, 12:11:59  
Job time : 951.552 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:24:34 ; Search time 256.262 Seconds  
(without alignments)  
1390.474 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQSQGGGFAIPI.....QTKSGGTRTGNVLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO\_spool\_p/US09684215/runat\_05092003.072201.548/app\_query.fasta\_1.853  
-DB=N\_Geneseq\_19Jun03 -QWTF=fastap -SUFFIX=ring -MINWATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215\_CGN\_1\_1\_626\_@runat\_05092003.072201.548 -NCPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	675	100.0	396	23	AAL40769	Nucleotide sequenc
2	675	100.0	672	23	AAL40772	Nucleotide sequenc
3	675	100.0	702	20	AA220206	Mycobacterium tube
4	675	100.0	702	23	AAL40770	Nucleotide sequenc
5	675	100.0	702	24	ABK14140	DNA encoding anti
6	675	100.0	1002	24	ABK14077	Mycobacterium tube
7	675	100.0	1002	24	AAD47078	Mycobacterium tube
8	675	100.0	1002	24	AAD28336	Mycobacterium sp.
9	675	100.0	1002	24	AAD28337	Mycobacterium spec
10	675	100.0	1068	20	AA334251	Mycobacterium spec
11	675	100.0	1143	20	AA334252	Mycobacterium spec
12	675	100.0	1629	22	AA303793	M. tuberculosis DN
13	675	100.0	1742	23	AAL40771	Nucleotide sequenc
14	675	100.0	1871	23	AAL40768	Nucleotide sequenc
15	675	100.0	2190	24	AAD47084	Mycobacterium sp.
16	675	100.0	2190	24	AAD28343	Mycobacterium spec
17	675	100.0	2191	23	AAL40773	Nucleotide sequenc
18	675	100.0	2286	24	ABK14128	DNA encoding anti
19	675	100.0	2287	20	AA220194	Mycobacterium tube
20	675	100.0	2287	24	AAD47083	Mycobacterium sp.
21	675	100.0	2287	24	AAD28342	Mycobacterium spec
22	675	100.0	2808	24	AAD47110	Mycobacterium sp.
23	675	100.0	4403765	22	AA199683	Mycobacterium tube
24	675	100.0	4411529	22	AA199682	Mycobacterium tube
25	670	99.3	447	18	AAT91466	Mycobacterium tube
26	670	99.3	447	18	AAT91403	M. tuberculosis im
27	670	99.3	447	19	AAV64450	Mycobacterium tube
28	670	99.3	447	19	AAV44342	M. tuberculosis an
29	670	99.3	447	20	AZ19252	M. tuberculosis re
30	670	99.3	447	20	AZ19040	M. tuberculosis re
31	670	99.3	447	22	AA303780	M. tuberculosis DN
32	670	99.3	447	24	AAD47080	Mycobacterium tube
33	670	99.3	447	24	AAD28339	Mycobacterium spec
34	670	99.3	1872	18	AAT91477	Mycobacterium tube
35	670	99.3	1872	18	AAT91414	Mycobacterium tube
36	670	99.3	1872	19	AAV64463	M. tuberculosis im
37	670	99.3	1872	19	AAV44355	Mycobacterium tube
38	670	99.3	1872	20	AZ19265	M. tuberculosis an
39	670	99.3	1872	20	AZ19053	M. tuberculosis re
40	670	99.3	1872	22	AA303781	M. tuberculosis DN
41	670	99.3	1872	24	AAD47076	Mycobacterium tube
42	670	99.3	1872	24	AAD28335	Mycobacterium spec
43	646	95.7	1067	22	AAH51961	Mycobacterium tube
44	632	93.6	675	22	AAH54132	Human /M. tubercul
45	632	93.6	675	22	AAH93896	Ral2-P510S-C const

ALIGNMENTS

RESULT 1

AAL40769  
ID AAL40769 standard; DNA; 396 BP.

AC AAL40769;

DT 03-OCT-2002 (first entry)

DE Nucleotide sequence encoding Ral2 protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.

OS Unidentified.

XX Key Location/Qualifiers

FH

```

CDS      1..396
          /tag= a
          /product= "Ral2 protein"
          /note= "No start or stop codon"
WO200125401-A2.
12-APR-2001.
06-OCT-2000; 2000WO-US27652.
07-OCT-1999; 99US-0158585.
(CORI-) CORIXA CORP.
Skeiky Y, Guderian J;
WPI: 2001-266299/27.
P-PSDB; AAO22138.
Recombinant nucleic acid molecule for producing high yield expression
of desired fusion polypeptides, encodes fusion polypeptide comprising
Mycobacterium tuberculosis coding sequence and heterologous polypeptide
-
Claim 1; Fig 2; 39pp; English.
The invention relates to a recombinant nucleic acid molecule encoding a
fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
kDa C-terminal fragment of serine protease antigen MTB32A of
Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
The recombinant fusion nucleic acids and polypeptides are useful for
providing stable and high yield expression of fusion polypeptides of both
eukaryotic and prokaryotic origin and to encode a protein product for use
as an antigen for detecting serum antibodies. The presence of serum
antibodies to M. tuberculosis antigens in an individual indicates that
the individual is infected with it. The fusion polypeptides are useful as
sources of proteins for monitoring binding of serum antibodies to fusion
proteins and as an immunogen to induce and/or enhance immune responses.
The coding sequences can be ligated with a coding sequence of another
molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
can be used in vivo as a DNA vaccine. This polynucleotide sequence
represents the DNA encoding the Ral2 protein.
Sequence 396 BP; 68 A; 132 C; 134 G; 52 T; 0 other;

Alignment Scores:
Pred. No.:      Length:      396
Score:          675.00      Matches:    132
Percent Similarity: 100.00%   Conservatives: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     100.00%   Indels:      0
DB:              23        Gaps:         0

US-09-684-215A-4 (1-132) x AAL40769 (1-396)
QY      1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB      1 ACGGGCGGCTCGGATAACATCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60
QY      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
DB      61 GGGCAGCGCATGGCGATCCGGGGCCAGATCCATCGGGGGGGTCAACCCACCGTTTCAT 120
QY      41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
DB      121 ATCGGGCCPACCGCTTCCTCGCTGGGTGTGTGTCGACAAACAGCGCAACGGCGCAGCA 180
QY      61 ValGlnArgValValGlySerIleProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB      181 GTCCAAACGGGTGGTGGGAGCGCTCCGGGGCAAGTCTCGGCATCTCCACCGGGCGACGTG 240
QY      81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db      241 ATCACCGGGTGGAGGGCGCTCCGATCACTCGGCACCGCATGGCGACGCGCTTAAC 300
QY      101 GlyHisHisPzGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db      301 GGGCATCATCGCGGTGACGTCTATCTCGGTGACCTTGGCAACCAAGTCGGGGCGGCGCGCT 360
QY      121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db      361 ACAGGGAACGTGACATTGGCCGAGGAGGCCCGCCGCGC 396

RESULT 2
AAL40772
ID      AAL40772 standard; DNA; 672 BP.
XX
AC      AAL40772;
XX
DT      03-OCT-2002 (first entry)
XX
DE      Nucleotide sequence encoding Ral2-mammaglobin fusion protein.
KW      Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW      vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW      immunogen; cytokine; gene; ds.
XX
OS      Chimeric - Mammalian.
XX      Chimeric - Unidentified.
XX
FH      Key Location/Qualifiers
FT      CDS 4..666 /tag= a /product= "Ral2-mammaglobin fusion protein"
FT
XX
PN      WO200125401-A2.
XX
PD      12-APR-2001.
XX
PF      06-OCT-2000; 2000WO-US27652.
XX
PR      07-OCT-1999; 99US-0158585.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Skeiky Y, Guderian J;
XX
DR      WPI; 2001-266299/27.
XX      P-PSDB; AAO22141.
XX
PT      Recombinant nucleic acid molecule for producing high yield expression
PT      of desired fusion polypeptides, encodes fusion polypeptide comprising
PT      Mycobacterium tuberculosis coding sequence and heterologous polypeptide
PT
XX
PS      Disclosure; Fig 5; 39pp; English.
XX
CC      The invention relates to a recombinant nucleic acid molecule encoding a
CC      fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
CC      kDa C-terminal fragment of serine protease antigen MTB32A of
CC      Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC      The recombinant fusion nucleic acids and polypeptides are useful for
CC      providing stable and high yield expression of fusion polypeptides of both
CC      eukaryotic and prokaryotic origin and to encode a protein product for use
CC      as an antigen for detecting serum antibodies. The presence of serum
CC      antibodies to M. tuberculosis antigens in an individual indicates that
CC      the individual is infected with it. The fusion polypeptides are useful as
CC      sources of proteins for monitoring binding of serum antibodies to fusion
CC      proteins and as an immunogen to induce and/or enhance immune responses.
CC      The coding sequences can be ligated with a coding sequence of another
CC      molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC      can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC      represents the DNA encoding the Ral2 protein.
XX
SQ      Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 other;
```

Alignment Scores:		4.96e-56	Length: 672
Pred. No.:	Score:	675.00	Matches: 132
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		23	Gaps: 0
US-09-684-215A-4 (1-132) x AAL40772 (1-672)			
QY	1	ThralaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
DB	25	ACGGCCCGTCCGATACCTCCAGCTGTCAGGGTGGCAGGATTCCGATTCGGATC	84
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
DB	85	GGCAGCGATGGCGATCGCGGCAGATCGATCGGTGGGGGTCAACCCACCGTTCAT	144
QY	41	IleGlyProThrAlaPheLeuGlyLeuValValAspAsnAsnGlyAlaArg	60
DB	145	ATCGGGCTACCGCTTCCTCGGCTTGGTGTTCGACAAACAGCGGCGCACGA	204
QY	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
DB	205	GTCCACGCGTGTGGGAGCGCTCCGCGGCAAGTCTCCACCGCGGCGGTG	264
QY	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
DB	265	ATCACCGCGTGCAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCTTAAC	324
QY	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg	120
DB	325	GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCTGGGCGGCACGCGT	384
QY	121	ThrGlyAsnValThrLeuAlaGluGlyProProIle	132
DB	385	ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC	420
RESULT 4			
AAZ20206	ID	AAZ20206 standard; DNA; 702 BP.	
XX	AC	AAZ20206;	
XX	DT	17-JAN-2000 (first entry)	
XX	DE	Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.	
XX	KW	Tuberculosis; antigen; fusion protein; Mtb24; Ral2; DPPD;	
XX	KW	diagnosis; therapy; vaccine; immunogen; ss.	
XX	OS	Mycobacterium tuberculosis.	
XX	PN	W09951748-A2.	
XX	PD	14-OCT-1999.	
XX	PF	07-APR-1999; 99WO-US07717.	
XX	PR	07-APR-1998; 98US-0056556.	
XX	PR	30-DEC-1998; 98US-0223040.	
XX	PA	(CORI-) CORIXA CORP.	
XX	PI	Skeiky YAW, Alderson M, Campos-Neto A;	
XX	DR	WPT; 1999-601610/51.	
XX	DR	P-PSDB; AAY32071.	
XX	PT	New fusion proteins useful for diagnosis, prevention and treatment of tuberculosis -	

PS	XX	Example; Fig 13A-B; 83pp; English.	
CC	CC	This DNA sequence includes a coding region for a recombinant	
CC	CC	Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),	
CC	CC	termed Mtb24, composed of the antigens Ral2 and DPPD. The	
CC	CC	DNA is useful for the recombinant production of the fusion protein.	
CC	CC	Coding sequences for the antigens were modified by PCR in order	
CC	CC	to facilitate their fusion and subsequent expression of the fusion	
CC	CC	protein, and then ligated. The invention provides fusion proteins	
CC	CC	(see AAY32059-71) containing at least 2 M. tuberculosis antigens. The	
CC	CC	new fusion proteins and polynucleotides encoding them are useful as	
CC	CC	vaccines for preventing tuberculosis (claimed), for diagnosis (via	
CC	CC	in vitro assays or intradermal skin tests for detection of anti-M.	
CC	CC	tuberculosis antibodies), monitoring of disease progression, and	
CC	CC	treatment of tuberculosis. They are more effective immunogens than	
XX	XX	mixtures of the individual protein components.	
SQ	SQ	Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;	

Alignment Scores:		5.22e-56	Length: 702
Pred. No.:	Score:	675.00	Matches: 132
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		20	Gaps: 0
US-09-684-215A-4 (1-132) x AAZ20206 (1-702)			
QY	1	ThralaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
DB	25	ACGGCCCGTCCGATACCTCCAGCTGTCAGGGTGGCAGGATTCCGATTCGGATC	84
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
DB	85	GGCAGCGATGGCGATCGCGGCAGATCGATCGGTGGGGGTCAACCCACCGTTCAT	144
QY	41	IleGlyProThrAlaPheLeuGlyLeuValValAspAsnAsnGlyAlaArg	60
DB	145	ATCGGGCTACCGCTTCCTCGGCTTGGTGTTCGACAAACAGCGGCGCACGA	204
QY	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
DB	205	GTCCACGCGTGTGGGAGCGCTCCGCGGCAAGTCTCCACCGCGGCGGTG	264
QY	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
DB	265	ATCACCGCGTGCAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCTTAAC	324
QY	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg	120
DB	325	GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCTGGGCGGCACGCGT	384
QY	121	ThrGlyAsnValThrLeuAlaGluGlyProProIle	132
DB	385	ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC	420
RESULT 4			
AAZ40770	ID	AAZ40770 standard; DNA; 702 BP.	
XX	XX	AAZ40770;	
XX	DT	03-OCT-2002 (first entry)	
XX	DE	Nucleotide sequence encoding Ral2-DPPD fusion protein.	
XX	KW	Ral2; serine protease antigen; Mtb32A; Mycobacterium tuberculosis;	
XX	KW	vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;	
XX	KW	immunogen; cytokine; gene; ds.	
XX	OS	Chimeric - Unidentified.	

Key Location/Qualifiers  
 CDS 4..696  
 /tag= a  
 /product= "Ra12-DPPD fusion protein"

WO200125401-A2.

12-APR-2001.

06-OCT-2000; 2000WO-US27652.

07-OCT-1999; 99US-0158585.

(CORI-) CORIXA CORP.

Skelky Y, Guderian J;

WPI: 2001-266299/27.

P-PSDB; AAO22139.

Recombinant nucleic acid molecule for producing high yield expression of desired fusion polypeptides, encodes fusion polypeptide comprising Mycobacterium tuberculosis coding sequence and heterologous polypeptide

Example 1; Fig 3; 39pp; English.

The invention relates to a recombinant nucleic acid molecule encoding a fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14 kDa C-terminal fragment of serine protease antigen MB32A of Mycobacterium tuberculosis, and a heterologous polynucleotide sequence. The recombinant fusion nucleic acids and polypeptides are useful for providing stable and high yield expression of fusion polypeptides of both eukaryotic and prokaryotic origin and to encode a protein product for use as an antigen for detecting serum antibodies. The presence of serum antibodies to M. tuberculosis antigens in an individual indicates that the individual is infected with it. The fusion polypeptides are useful as sources of proteins for monitoring binding of serum antibodies to fusion proteins and as an immunogen to induce and/or enhance immune responses. The coding sequences can be ligated with a coding sequence of another molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and can be used in vivo as a DNA vaccine. This polynucleotide sequence represents the DNA encoding the Ra12-DPPD fusion protein.

SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

Alignment Scores:

Pred. No.: 5,22e-56 Length: 702  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 23 Gaps: 0

US-09-684-215A-4 (1-132) x AAL40770 (1-702)

QY 1 ThrAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 DB 25 ACGGCGCGGTCGGAATCCAGCTGCCAGGTCGAGGATTCGCCATTCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 DB 85 GGGCAGGCGGTCGCGATCGCGGCGAGATCCGATCGGGTGGGTCACCCACCGTTCAT 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60  
 DB 145 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTTCGACACACACGCGACGCGCAGCA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCCAAACGCTGGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGCACGTG 264  
 QY 81 IleThrAlaValAspGlyAlaPheAlaProfileAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

DB 265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCACCGGATGGCGGACGCGTTAAC 324  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 325 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGCGGGGCGACGGGT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGACGTCGACATTGGCCGAGGACCCCGGCC 420  
 RESULT 5  
 ABK14140  
 ID ABK14140 standard; DNA; 702 BP.  
 XX  
 AC ABK14140;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE DNA encoding antigenic fusion protein Ra12-DPPD (Mtb24).  
 KW: Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;  
 KW: tuberculostatic; immunogen; vaccine; Ra12-DPPD; Mtb24.  
 XX  
 OS Chimeric - Mycobacterium tuberculosis.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..702  
 /tag= a  
 /product= "Mtb24 #1"  
 /partial  
 /transl\_except= (pos:694..696, aa:Xaa)  
 /note= "No start or stop codon. Xaa= In frame stop codon"  
 CDS 2..702  
 /tag= b  
 /product= "Mtb24 #2"  
 /partial  
 /note= "No start or stop codon"  
 /transl\_except= (pos:263..265, aa:Xaa)  
 /transl\_except= (pos:353..355, aa:Xaa)  
 /transl\_except= (pos:395..397, aa:Xaa)  
 /transl\_except= (pos:470..472, aa:Xaa)  
 /transl\_except= (pos:701..702, aa:Ser)  
 /note= "This codon has an apparent 1 nucleotide deletion which alters the reading frame."  
 Xaa= In frame stop codon"  
 CDS 3..701  
 /tag= c  
 /product= "Mtb24 #3"  
 /partial  
 /note= "No start or stop codon"  
 /transl\_except= (pos:1..2, aa:Pro)  
 /transl\_except= (pos:39..41, aa:Xaa)  
 /transl\_except= (pos:321..323, aa:Xaa)  
 /transl\_except= (pos:339..341, aa:Xaa)  
 /transl\_except= (pos:450..452, aa:Xaa)  
 /transl\_except= (pos:621..623, aa:Xaa)  
 /note= "No start or stop codon. Xaa= In frame stop codon"  
 XX  
 US2002009459-A1.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 07-APR-1999; 99US-0287849.  
 XX  
 PR 13-MAR-1997; 97US-0818112.  
 PR 01-OCT-1997; 97US-0942578.  
 PR 18-FEB-1998; 98US-0925197.  
 PR 07-APR-1998; 98US-0956556.  
 PR 30-DEC-1998; 98US-0223040.  
 XX  
 PA (REED/) REED S G.  
 PA (SKEI/) SKEIKY Y A.



PA (DILL/) DILLON D C.  
PA (ALDE/) ALDERSON M.  
XX (CAMP/) CAMPOS-NETO A.  
XX  
XX Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
XX  
XX WPI: 2002-171134/22.  
DR P-PSDB: AAU74600, AAU76541, AAU76542.  
XX  
XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for  
PT diagnosing, treating or preventing M. tuberculosis infection,  
PT particularly as vaccine for treating or preventing tuberculosis  
XX  
XX Examples; Fig 13; 62pp; English.  
XX  
XX The invention relates to a purified polypeptide which induces an immune  
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are  
CC useful for diagnosing, treating or preventing M. tuberculosis infection,  
CC particularly tuberculosis infection. In particular, the polypeptides are  
CC useful as a vaccine formulation with an adjuvant to afford long-term  
CC protection in animals against the development of tuberculosis. The  
CC protein coding sequence may be used to encode a protein product for use  
CC as an immunogen to induce and/or enhance an immune response to M.  
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis  
CC fusion protein of the invention. This polynucleotide encodes 3 different  
XX proteins, each in a different reading frame.  
XX  
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

## Alignment Scores:

Pred. No.: 5,22e-56 Length: 702  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x ABK14140 (1-702)

QY 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 25 ACGGCGCGTCCGATACCTCCAGCTGTCAGGGTGGCAGGATCCCATTCGATC 84  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
DB 85 GGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGTCAACCACCGTTCAT 144  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60  
DB 145 ATCGGGCCCTACCGCCCTCCTCGGCTTGGGTGTGTCGACAAACACGCGCACCA 204  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 205 GTCCACGCGTGTCTGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 265 ATCACCGCGTGCAGCGCTCGATCAACTCGGCCACCGCGATGGCGACGCGCTAAC 324  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg 120  
DB 325 GGCATCATCCCGGTGACGTCAFTCTGGTGACCTGGCAACCAAGTGGGGCGCGCGGT 384  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 385 ACAGGAAACGTGACATTGGCCGAGGAGGCCCGCCGCC 420

## RESULT 6

AA047077  
ID AA047077 standard; DNA: 1002 BP.  
XX  
AC AA047077;  
XX

DT 27-JAN-2003 (first entry)  
XX  
XX Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.  
DE  
XX Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen;  
KW gene; ds.  
XX  
XX Mycobacterium tuberculosis.  
OS  
XX  
XX Key Location/Qualifiers  
CDS 4..996  
FT /tag= a  
FT /product= "Ra35 mature antigenic protein"  
FT /transl\_except= (pos:547..549, aa:Asp)  
FT /transl\_except= (pos:550..552, aa:Ser)  
XX  
XX WO200272792-A2.  
XX  
XX 19-SEP-2002.  
XX  
XX 13-MAR-2002; 2002WO-US08223.  
XX  
XX 13-MAR-2001; 2001US-275837P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Brannon M, Guderian J;  
XX  
XX WPI: 2002-759844/82.  
DR P-PSDB: AAE29702.  
XX  
XX New recombinant nucleic acid molecule comprising a Leishmania TSA,  
PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective  
PT immunity against pathogenic microorganisms e.g. Leishmania and  
PT Mycobacterium tuberculosis  
XX  
XX Disclosure; Page 79; 155pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
CC are used in methods for eliciting immune response in mammals. They are  
CC useful as vaccines to elicit protective immunity against pathogenic  
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
CC polypeptides are used for enhancing the expression of polynucleotides,  
CC as in vivo diagnostic agents and for raising antibodies in a non-human  
CC animal. The invention is used in gene therapy. The present sequence is  
CC M. tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen  
CC encoding DNA.  
XX  
SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;

## Alignment Scores:

Pred. No.: 7.89e-56 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x AAD47077 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 598 ACGGCGCGTCCGATACCTCCAGGTGTCAGGGTGGCAGGATTCGCATTCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
DB 658 GGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTGTCACCCACCGTTCAT 717

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
DB 718 ATCGGGCTACCGCTTCTCGGCTGGGTGTGTCGACAAACAAGCGGCGCACGA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 778 GTCCAAACGGTGTGGGAGCGCTCCGGCGCAAGCTCGGCATCTCCACGGCGACGTG 837  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 838 ATCACCAGCGGTGCGAGCGCTCCGATCACTCCGCGCACCGGATGCGGACGGCTTAAC 897  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 898 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGCACGGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 958 ACAGGGAACGTGACATTGGCGAGGACCCCGGCC 993  
RESULT 7  
AAD47078  
ID AAD47078 standard; DNA; 1002 BP.  
XX  
AC AAD47078;  
XX  
XX  
DT 27-JAN-2003 (first entry)  
XX  
DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.  
XX  
XX Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;  
KW gene; antigen; ds.  
XX  
OS Mycobacterium tuberculosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT CDS 4..996  
FT /\*tag= a  
FT /product= "Ra35FLMutSA mutant antigenic protein"  
XX  
XX WO200272792-A2.  
PN  
XX  
PD 19-SEP-2002.  
XX  
PF 13-MAR-2002; 2002WO-US08223.  
XX  
PR 13-MAR-2001; 2001US-275837P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky Y, Brannon M, Guderian J;  
XX  
DR WPI: 2002-759844/82.  
DR P-PSDB; AAE29703.  
XX  
PT New recombinant nucleic acid molecule comprising a Leishmania TSA,  
PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective  
PT immunity against pathogenic microorganisms e.g. Leishmania and  
PT Mycobacterium tuberculosis  
XX  
XX Disclosure; Page 80-81; 155pp; English.  
XX  
CC The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
CC are used in methods for eliciting immune response in mammals. They are  
CC useful as vaccines to elicit protective immunity against pathogenic  
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
CC polypeptides are used for enhancing the expression of polynucleotides,  
CC

CC as in vivo diagnostic agents and for raising antibodies in a non-human  
CC animal. The invention is used in gene therapy. The present sequence is  
CC M. tuberculosis Ra35FLMutSA mutant antigenic protein encoding DNA.  
XX  
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;  
Alignment Scores:  
Pred. No.: 7.89e-56 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
US-09-684-215A-4 (1-132) x AAD47078 (1-1002)  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 598 ACGGCGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 658 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCTACCCACCGCTTCA 717  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
DB 718 ATCGGGCTACCGCTTCTCGGCTGGGTGTGTCGACAAACAAGCGGCGCACGA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 778 GTCCAAACGGTGTGGGAGCGCTCCGGCGCAAGCTCTCGGCATCTCCACGGCGACGTG 837  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 838 ATCACCAGCGGTGCGAGCGCTCCGATCACTCCGCGCACCGGATGCGGACGGCTTAAC 897  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 898 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGCACGGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 958 ACAGGGAACGTGACATTGGCGAGGACCCCGGCC 993  
RESULT 8  
AAD28336  
ID AAD28336 standard; cDNA; 1002 BP.  
XX  
AC AAD28336;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.  
XX  
KW Fusion protein; antigen; serological sensitivity; immune response;  
KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.  
XX  
OS Mycobacterium sp.  
FH Key Location/Qualifiers  
FT CDS 4..996  
FT /\*tag= a  
FT /product= "Ra35 mature protein"  
FT /transl\_except= (pos:547..549, aa:Asp)  
FT /transl\_except= (pos:550..552, aa:Ser)  
XX  
XX WO200198460-A2.  
XX  
PD 27-DEC-2001.  
XX  
XX 20-JUN-2001; 2001WO-US19959.  
XX  
PR 20-JUN-2000; 2000US-0597796.

PR 01-FEB-2001; 2001US-265737P.  
XX (CORI-) CORIXA CORP.  
XX Skeiky Y, Reed S, Alderson M;  
XX WPI; 2002-147798/19.  
DR P-PSDB; AAE17566.  
XX  
XX Composition comprising MTB39 antigen and MTB32A antigen from  
PT Mycobacterium species, useful for eliciting immune response in a  
PT subject  
XX  
XX Disclosure; Page 95; 136pp; English.  
XX  
XX The present invention relates to fusion proteins containing at least  
CC two Mycobacterium species antigens, nucleotides encoding them and  
CC compositions comprising such fusion proteins. The present invention  
CC particularly relates to nucleic acids encoding fusion proteins that  
CC include two or more individual M. tuberculosis antigens which increase  
CC the serological sensitivity of sera from individuals infected with  
CC tuberculosis and methods for their use in diagnosis, prevention and  
CC treatment of tuberculosis infection. Sequences of the invention are  
CC useful for eliciting an immune response in a mammal, e.g., human,  
CC immunised with BCG. They are useful in the diagnosis, treatment and  
CC prevention of Mycobacterium infection. The fusion proteins and the  
CC polynucleotides are useful as diagnostic tools in patients infected  
CC with Mycobacterium, in vitro and in vivo assays for detecting humoral  
CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
CC diagnosis of an infection or monitoring of disease progression, as  
CC immunogens to generate or elicit a protective immune response in a  
CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
CC animal. Sequences of the invention are also used as vaccines. MTB32A  
CC fusion proteins of the invention are useful as in vivo diagnostic agents  
CC for intradermal skin test. The present sequence is a cDNA encoding  
CC Mycobacterium species MTB32A (Ra32FL) mature protein.  
XX  
XX Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;  
SQ

Alignment Scores:  
Pred. No.: 7,89e-56 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x AAD28336 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 598 ACGGCGCGTCCGATAACTTCCAGCTGCCAGGGTGGCGAGGATTCGCCATTCCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 658 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 717  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60  
DB 718 ATCGGGCTACCGCTTCCTCGGTGGGTGGTGTGCACACACACGCGCGCACGA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 778 GTCCACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGGCGGTG 837  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 838 ATCACCAGCGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAC 897  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArg 120  
DB 898 GGGCATCATCCCGTGGTACGTCATCTCGTGACCTGGCAACCAAGTGGCGGCGGCGCGT 957

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 958 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 993  
RESULT 9  
AAD28337  
ID AAD28337 standard; cDNA; 1002 BP.  
XX  
XX AAD28337;  
XX  
XX 22-APR-2002 (first entry)  
XX  
XX Mycobacterium species Ra35FLMutSA mutant cDNA.  
XX  
XX Fusion protein; antigen; serological sensitivity; immune response;  
XX tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.  
XX  
XX Mycobacterium sp.  
XX  
XX Key Location/Qualifiers  
XX CDS 4..996 /\*tag= a  
XX /product= "Ra35FLMutSA protein"  
XX  
XX WO200198460-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 20-JUN-2001; 2001WO-US19959.  
XX  
XX 20-JUN-2000; 2000US-0597796.  
XX  
XX 01-FEB-2001; 2001US-265737P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Reed S, Alderson M;  
XX  
XX WPI; 2002-147798/19.  
XX  
XX P-PSDB; AAE17567.  
XX  
XX Composition comprising MTB39 antigen and MTB32A antigen from  
XX Mycobacterium species, useful for eliciting immune response in a  
XX subject  
XX  
XX Claim 67; Page 96-97; 136pp; English.  
XX  
XX The present invention relates to fusion proteins containing at least  
XX two Mycobacterium species antigens, nucleotides encoding them and  
XX compositions comprising such fusion proteins. The present invention  
XX particularly relates to nucleic acids encoding fusion proteins that  
XX include two or more individual M. tuberculosis antigens which increase  
XX the serological sensitivity of sera from individuals infected with  
XX tuberculosis and methods for their use in diagnosis, prevention and  
XX treatment of tuberculosis infection. Sequences of the invention are  
XX useful for eliciting an immune response in a mammal, e.g., human,  
XX immunised with BCG. They are useful in the diagnosis, treatment and  
XX prevention of Mycobacterium infection. The fusion proteins and the  
XX polynucleotides are useful as diagnostic tools in patients infected  
XX with Mycobacterium, in vitro and in vivo assays for detecting humoral  
XX antibodies or cell-mediated immunity against M. tuberculosis, for the  
XX diagnosis of an infection or monitoring of disease progression, as  
XX immunogens to generate or elicit a protective immune response in a  
XX patient and for raising anti-M. tuberculosis antibodies in a non-human  
XX animal. Sequences of the invention are also used as vaccines. MTB32A  
XX fusion proteins of the invention are useful as in vivo diagnostic agents  
XX for intradermal skin test. The present sequence is a cDNA encoding  
XX Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA.  
XX  
XX Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;  
SQ

Alignment Scores:  
Pred. No.: 7,89e-56 Length: 1002  
Score: 675.00 Matches: 132



PI Gicquel B, Lim EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y;  
 PI Guigueno A;  
 XX WPI; 1999-181045/15.  
 DR P-PSDB; AAY04830.  
 XX Mycobacterial DNA vectors containing reporter constructs - for  
 PT identifying coding or promoter sequences involved in  
 PT infection-associated protein expression  
 XX Claim 22; Fig 50F; 309pp; French.  
 XX Sequences AAX34001-X34252 represent nucleic acids encoding secreted  
 CC proteins from various Mycobacterium species-microorganisms. The  
 CC nucleotide sequences can be used as primers and probes for methods  
 CC for detecting and identifying mycobacteria, especially belonging to  
 CC the M. tuberculosis complex. The encoded proteins can be used in  
 CC vaccines for immunisation against a bacterial or viral infection.  
 XX Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 other;  
 SO Alignment Scores:  
 Pred. No.: 9.2e-56 Length: 1143  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-684-215A-4 (1-132) x AAX34252 (1-1143)

Qy 1 ThrAlaLaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 745 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGTTGGCAGGATTCGGCATTCGCATC 804

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 805 GGGCAGGCGATGGCGATGGCGGCGCAGATCGGTTGGGGTGGGGTGGGGTGGGGTGGGGT 864

Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60  
 Db 865 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACACACACGCGGCGCAGCA 924

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 925 GTCCACGCGTGTGCGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGCGTG 984

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 985 ATCACCAGCGGTGACGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 1044

Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 1045 GGGCATCATCCGCTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGCGCGT 1104

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 1105 ACAGGGAACGTGACATTTGGCGGAGGAGGCCCGCGGCC 1140

RESULT 12  
 AAS03793  
 ID AAS03793 standard; DNA; 1629 BP.  
 XX AAS03793;  
 AC AAS03793;  
 XX 29-AUG-2001 (first entry)  
 XX M. tuberculosis DNA encoding antigen HTCC#1 fusion protein #4.  
 DE TBRA12-HTCC#1; antigen; vaccine; tuberculosis;  
 KW AIDS; acquired immunodeficiency disease; His Tag; ds.  
 XX Mycobacterium tuberculosis.  
 OS

## Synthetic.

OS Key Location/Qualifiers  
 XX CDS 1..1629  
 FT /\*tag= a  
 FT /product= "TBRA12-HTCC#1"  
 FT /transl\_except= (pos:1621..1623,aa:Xaa)  
 FT /note= "Xaa= In frame STOP codon"  
 FT /partial  
 FT /note= "No start or stop codon. Although the sequence  
 FT does contain an in frame stop codon, 2 further amino  
 FT acids are shown in Figure 8 as being encoded by the  
 FT present sequence, without a further stop codon".  
 FT 25..426  
 FT misc\_feature /\*tag= b  
 FT /note= "Region derived from TBRA12"  
 FT 427..444 C  
 FT misc\_feature /\*tag= c  
 FT /note= "Region derived from Thrombin"  
 FT 445..1629  
 FT misc\_feature /\*tag= d  
 FT /note= "Region derived from HTCC#1"

WO200124820-A1.

12-APR-2001.

10-OCT-2000; 2000WO-US28095.

07-OCT-1999; 99US-0158338.

07-OCT-1999; 99US-0158425.

(CORI-) CORIXA CORP.

Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;

WPI; 2001-290576/30.

Vaccinating against Mycobacteria infections in mammals using fusion  
 proteins comprising combinations of heterologous antigens

Example 2; Fig 8; 168pp; English.

The sequence encodes Mycobacterium tuberculosis fusion protein,  
 TBRA12-HTCC#1 and includes a His tag at the N-terminus to aid  
 purification. Compositions comprising at least 2 heterologous antigens,  
 as a fusion protein, and vectors expressing the fusion proteins are  
 used as vaccines to prophylactically immunise mammals (especially  
 humans) against infection by Mycobacteria. The compositions  
 contain at least 2 heterologous antigens that increase the serological  
 sensitivity of individuals infected with tuberculosis, a disease  
 frequently affecting patients with acquired immunodeficiency disease,  
 AIDS.

SQ Sequence 1629 BP; 298 A; 500 C; 535 G; 296 T; 0 other;

## Alignment Scores:

Pred. No.: 1.39e-55 Length: 1629  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-684-215A-4 (1-132) x AAS03793 (1-1629)

Qy 1 ThrAlaLaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 25 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGTTGGCAGGATTCGGCATTCGCATC 84

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 85 GGGCAGGCGATGGCGATGGCGGCGCAGATCCGATCGGGTGGGGGTGGGGTGGGGTGGGGT 144

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
 DB 145 ATCGGCGCTACCGCTTCTCGGCTGGGTGTGTCGACAAACGCGACGCGACGA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCCAAACGCGTGGTGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGACGCTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 265 ATCAGCGCGTGCAGCGCGCTCGGATCACTCGGCACCGCGATGGCGGCGGCTTAAAC 324  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 325 GGGCATCATCCCGTGACGTATCTCCGTGACCTGGCAAAACCAAGTCGGCGGCGACGCT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 420

RESULT 13  
 AAL40771  
 ID AAL40771 standard; DNA; 1742 BP.  
 XX  
 AC AAL40771;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Nucleotide sequence encoding Ral2-WT1 fusion protein.  
 XX  
 KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
 KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
 KW immunogen; cytokine; gene; ds.  
 XX  
 OS Chimeric - Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 4..1740  
 FT /\*tag= a  
 FT /product= "Ral2-WT1 fusion protein"  
 XX  
 PN WO200125401-A2.  
 XX  
 PD 12-APR-2001.  
 XX  
 PF 06-OCT-2000; 2000WO-US27652.  
 XX  
 PR 07-OCT-1999; 99US-0158585.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skelky Y, Guderian J;  
 XX  
 DR WPI: 2001-266299/27.  
 DR P-PSDB; AA022140.  
 XX  
 PT Recombinant nucleic acid molecule for producing high yield expression  
 PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
 PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide  
 XX  
 PS Disclosure; Fig 4; 39pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that

the individual is infected with it. The fusion polypeptides are useful as  
 sources of proteins for monitoring binding of serum antibodies to fusion  
 proteins and as an immunogen to induce and/or enhance immune responses.  
 The coding sequences can be ligated with a coding sequence of another  
 molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 represents the DNA encoding the Ral2-WT1 fusion protein.  
 Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.5e-55 Length: 1742  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 23 Gaps: 0

US-09-684-215A-4 (1-132) x AAL40771 (1-1742)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 DB 25 ACAGCGCGCTCGGATCACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 DB 85 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTTCACCCACCGTTCAT 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
 DB 145 ATCGGCGCTACCGCTTCTCGGCTGGGTGTGTCGACAAACGCGACGCGCGACGA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCCAAACGCGTGGTGGGAGCGCTCGGCGGCAAGTCTCGGCATCTCCACCGCGACGCTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 265 ATCAGCGCGTGCAGCGCGCTCGGATCACTCGGCACCGCGATGGCGGCGGCTTAAAC 324  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 325 GGGCATCATCCCGTGACGTATCTCCGTGACCTGGCAAAACCAAGTCGGCGGCGACGCT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 420

RESULT 14  
 AAL40768  
 ID AAL40768 standard; DNA; 1871 BP.  
 XX  
 AC AAL40768;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Nucleotide sequence encoding MTB32A protein.  
 XX  
 KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
 KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
 KW immunogen; cytokine; gene; ds.  
 XX  
 OS Mycobacterium tuberculosis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 89..1156  
 FT /\*tag= a  
 FT /product= "Bacillus mycobacterium MTB32A protein"  
 XX  
 PN WO200125401-A2.  
 XX  
 PD 12-APR-2001.

PF 06-OCT-2000; 2000WO-US27652.  
PR 07-OCT-1999; 99US-0158585.  
XX (CORI-) CORIXA CORP.  
PA Skeiky Y, Guderian J;  
PI WPI; 2001-266299/27.  
DR P-PSDB; AAO22137.  
XX  
XX Recombinant nucleic acid molecule for producing high yield expression  
XX of desired fusion polypeptides, encodes fusion polypeptide comprising  
XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide  
XX  
XX Disclosure; Fig 1; 39pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
XX fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
XX kDa C-terminal fragment of serine protease antigen MTB32A of  
XX Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
XX The recombinant fusion nucleic acids and polypeptides are useful for  
XX providing stable and high yield expression of fusion polypeptides of both  
XX eukaryotic and prokaryotic origin and to encode a protein product for use  
XX as an antigen for detecting serum antibodies. The presence of serum  
XX antibodies to M. tuberculosis antigens in an individual indicates that  
XX the individual is infected with it. The fusion polypeptides are useful as  
XX sources of proteins for monitoring binding of serum antibodies to fusion  
XX proteins and as an immunogen to induce and/or enhance immune responses.  
XX The coding sequences can be ligated with a coding sequence of another  
XX molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
XX can be used in vivo as a DNA vaccine. This polynucleotide sequence  
XX represents the DNA encoding the Mycobacterium tuberculosis MTB32A  
XX protein.  
XX  
XX SQ Sequence 1871 BP; 317 A; 617 C; 605 G; 332 T; 0 other;

Alignment Scores:  
Pred. No.: 1.63e-55 Length: 1871  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 23 Gaps: 0

US-09-684-215A-4 (1-132) x AAL40768 (1-1871)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
DB 758 ACGGCGCGGTCCGATAAATCCAGCTGCCAGGTGGCAGGGATTCCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGCATGGCGATCCGGGGCCAAATCCATCCGTTGGGGGTCCACCCACCGTTTAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAlaArg 60  
DB 878 ATCGGCGCTACCGCTCTCTCGCTGGTGTGTGTCGACAAACGCGACGCGACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCACCGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 998 ATCACCGCGTCCGAGCGCTCCGATCACTCGGCACCGCATGGCGGACCGCTTAC 1057  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 1058 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACCAAGTCGGCGCGACGCT 1117  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

DB 1118 ACAGGGAACGTGACATTGGCCGAGGAGCCCGGCC 1153  
RESULT 15  
AAD47084  
ID AAD47084 standard; DNA; 2190 BP.  
XX  
XX AAD47084;  
XX  
XX 27-JAN-2003 (first entry)  
XX  
XX Mycobacterium sp. MTB72FMutSA fusion protein encoding DNA.  
XX  
XX Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen;  
XX Ra35MutSA; Ra12; MTB72MutSA; chimeric; gene; ds.  
XX  
XX Chimeric - Mycobacterium sp.  
XX Chimeric - Mycobacterium tuberculosis.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..2190  
XX /tag= a  
XX /product= "MTB72FMutSA fusion protein"  
XX  
XX WO200272792-A2.  
XX  
XX 19-SEP-2002.  
XX  
XX 13-MAR-2002; 2002WO-US08223.  
XX  
XX 13-MAR-2001; 2001US-275837P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Brannon M, Guderian J;  
XX  
XX WPI; 2002-759844/82.  
XX P-PSDB; AAE29709.  
XX  
XX New recombinant nucleic acid molecule comprising a Leishmania TSA,  
XX LeIF, M15 or 6F polynucleotide, useful as vaccine to elicit protective  
XX immunity against pathogenic microorganisms e.g. Leishmania and  
XX Mycobacterium tuberculosis  
XX  
XX Disclosure; Page 92-93; 155pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
XX fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
XX polynucleotide sequence encoding an antigen or an antigenic fragment from  
XX Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
XX polypeptide or its fragment. The Leishmania polynucleotide is selected  
XX from TSA, LeIF, M15, and 6F polynucleotides. Sequences of the invention  
XX are used in methods for eliciting immune response in mammals. They are  
XX useful as vaccines to elicit protective immunity against pathogenic  
XX microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
XX polypeptides are used for enhancing the expression of polynucleotides,  
XX as in vivo diagnostic agents and for raising antibodies in a non-human  
XX animal. The invention is used in gene therapy. The present sequence is  
XX a DNA encoding MTB72F fusion protein. This fusion protein comprises  
XX Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and  
XX TbH9 protein from Mycobacterium sp.  
XX  
XX SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 other;

Alignment Scores:  
Pred. No.: 1.96e-55 Length: 2190  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x AAD47084 (1-2190)

Qy	1	ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	22	ACGGCGGGTCCGATAACTTCACGCTGCCAGGTGGCAGGATTCGCATTCCGATC	81
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerClyGlySerProThrValHis	40
Db	82	GGGCAGGCGATGGCGATCCGGGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTCAT	141
Qy	41	IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg	60
Db	142	ATCGGGCCTACCGCTTCCTCGGGTCTGGGTGTGTCGACAAACACGGCAACGGCGCACGA	201
Qy	61	ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	202	GTCCAAACGCGTGTTCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGCAGCTG	261
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	262	ATCACCGGGGTTCAGCGGCTCCGATCAACTCGGCCACCGCATGGCGACGCGCTTAAC	321
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
Db	322	GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCACGCT	381
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132
Db	382	ACAGGGAACGTGACATTGGCCAGGGACCCCGGGCC	417

Search completed: September 5, 2003, 09:04:16  
Job time : 264.595 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:27:24 ; Search time 2280.41 Seconds  
(without alignments)  
1406.846 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGQGFPAIP.....QTKSGGTRGNTVLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_B/US09684215/runat\_05092003\_072202\_572/app\_query.fasta\_1.853  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215@cgn.1.1.4575 -runat\_05092003\_072202\_572 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estmu: \*  
4: em\_estov: \*  
5: em\_estpl: \*  
6: em\_estro: \*  
7: em\_estro: \*  
8: em\_estro: \*  
9: gb\_estl: \*  
10: gb\_estc: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mus: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rod: \*  
26: em\_gss\_phg: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*

29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	145.5	21.6	289	29	U82114	U82114 U82114 orde
2	126.5	18.7	1438	28	BH770798	BH770798 LLMGtag54
3	122	18.1	959	29	BZ549048	BZ549048 pacs1-60
4	120.5	17.9	726	14	CB679186	CB679186 OSJNEF02F
5	117.5	17.4	590	14	CB925602	CB925602 ABAL_22.F
6	114	16.9	603	28	AZ934428	AZ934428 BJ_Ba000
7	113	16.7	758	13	BQ514888	BQ514888 EST622303
8	106.5	15.8	1033	29	BZ561390	BZ561390 pacs2-164
9	103.5	15.3	719	28	AZ933900	AZ933900 BJ_Ba000
10	103.5	15.3	947	29	BZ549047	BZ549047 pacs1-60
11	103	15.3	726	28	AO989479	AO989479 Rfc00025
12	102	15.1	765	9	AJ558965	AJ558965 AJ558965
13	100.5	14.9	859	14	CA480614	CA480614 AGENCOURT
14	99	14.7	423	9	AI597611	AI597611 tn15f02.x
15	97.5	14.4	499	12	BI350520	BI350520 fr32b05.y
16	97.5	14.4	870	14	CA473840	CA473840 AGENCOURT
17	97	14.4	543	12	BI721127	BI721127 1031054B1
18	96	14.2	546	9	AW285510	AW285510 LG1_241_E
19	96	14.2	551	9	AW285527	AW285527 LG1_241_G
20	96	14.2	1093	13	BU557763	BU557763 AGENCOURT
21	96	14.2	1394	29	BZ576076	BZ576076 msh2_4761
22	95.5	14.1	677	12	BI378928	BI378928 BFLG1_000
23	95	14.1	909	14	CA975497	CA975497 AGENCOURT
24	94.5	14.0	728	12	BJ285991	BJ285991 BJ285991
25	94.5	14.0	1029	12	BI457360	BI457360 603185681
26	94	13.9	434	28	BHG38292	BHG38292 1008021F0
27	94	13.9	672	14	CD047378	CD047378 PSB027xJ
28	94	13.9	1321	13	BU509638	BU509638 AGENCOURT
29	93.5	13.9	600	29	CC345457	CC345457 OQAF70TH
30	93.5	13.9	907	14	CA471201	CA471201 AGENCOURT
31	93	13.8	689	12	BM448307	BM448307 DSA024D06
32	93	13.8	951	14	CA474469	CA474469 AGENCOURT
33	92.5	13.7	452	12	BM874057	BM874057 laa07a02.
34	92.5	13.7	860	14	CA471291	CA471291 AGENCOURT
35	92.5	13.7	1258	12	BM474560	BM474560 AGENCOURT
36	92	13.6	894	29	BZ578796	BZ578796 msh2_6009
37	92	13.6	899	14	CA471516	CA471516 AGENCOURT
38	91.5	13.6	513	12	BM036648	BM036648 fu0901.y
39	91.5	13.6	545	10	BE921575	BE921575 EST425344
40	91.5	13.6	548	13	BQ112264	BQ112264 EST597840
41	91.5	13.6	555	13	BQ479512	BQ479512 faa68g09.
42	91.5	13.6	570	14	CB364397	CB364397 2F001-P00
43	91.5	13.6	613	9	AW595047	AW595047 fk28e03.y
44	91.5	13.6	746	14	CD284690	CD284690 G39171.95
45	91.5	13.6	810	14	CA470510	CA470510 AGENCOURT

# ALIGNMENTS

RESULT 1

U82114

LOCUS

DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.

ACCESSION U82114

VERSION U82114.1

KEYWORDS GI:3647212

SOURCE GSS.

ORGANISM Mycobacterium leprae

Mycobacterium leprae

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;.

Corynebacterineae; Mycobacteriaceae; Mycobacterium.

REFERENCE 1 (bases 1 to 289)

U82114 289 bp DNA linear GSS 16-FEB-2001





```

QY 113 GlnThrLysSerGlyClyThrArgThrGlyAsnValThrLeu 126
Db 234 CAAGA-----GGCGCTGAATCTCTGGAGGTAACCTTG 202

RESULT 5
CB925602
LOCUS
DEFINITION
  CB925602 590 bp mRNA linear EST 28-APR-2003
  ABAL_22_F06_b1_A012 Absciscic acid-treated seedlings Sorghum bicolor
  cDNA clone ABAL_22_F06_A012 3', mRNA sequence.
ACCESSION
  CB925602
VERSION
  CB925602.1 GI:30161873
KEYWORDS
  EST.
SOURCE
  Sorghum bicolor (sorghum)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
  1 (bases 1 to 590)
  Cordonnier-Pratt M.-M., Wentzel V., Suzuki Y., Sugano S., Klein
  R.R., Liang C., Sun F., Sullivan R., Shah M., Buchanan C.D.,
  Eastman A. and Pratt L.H.
  An EST database from Sorghum: ABAL-treated seedlings
  Unpublished
  Other ESTs: ABAL_22_F06_g1_A012
  Contact: Cordonnier-Pratt MW
  Laboratory for Genomics and Bioinformatics
  The University of Georgia, Department of Plant Biology
  Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
  Tel: 706 542 1860
  Fax: 706 583 0210
  Email: mmpratt@uga.edu
  Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
  the Human Genome Center, University of Tokyo Institute of Medical
  Science; plant material and RNA prepared at Texas A & M University;
  sequencing done in the Laboratory for Genomics and Bioinformatics,
  University of Georgia. Sequence ends have been trimmed to exclude
  vector and regions below Phred quality 16. Three-prime sequences
  are presented as their reverse complement and have been trimmed to
  exclude polyA.
  Poly primer: Sug3 (CGACCTGCAGCTCGACACA)
  SEQ Y=Yes.
FEATURES
  source
    Location/Qualifiers
      1..590
      /organism="Sorghum bicolor"
      /mol_type="mRNA"
      /cultivar="IS3620C"
      /db_xref="taxon:4558"
      /clone="ABAL_22_F06_A012"
      /lab_host="DH10B-T1 phage-resistant E. coli"
      /clone_lib="Absciscic acid-treated seedlings"
      /note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
      library was prepared from polyA+ RNA from seedlings grown
      in hydroponic culture. After 12 days, medium was
      supplemented with 1 mM absciscic acid (ABA), while leaves
      were misted with a solution of 1 mM ABA. Roots and leaves
      were harvested after 3, 6, 12, and 24 hr and material from
      all time points was combined prior to RNA isolation.
      Double-stranded cDNA was cloned unidirectionally into
      different DraIII sites of the pME18S-FL3 vector (5'-prime
      DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG
      ). XhoI excises the cDNA insert."
BASE COUNT 136 a 137 c 156 g 161 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0509 Length: 590
Score: 117.50 Matches: 37
Percent Similarity: 44.19% Conservative: 20
Best Local Similarity: 28.68% Mismatches: 43
Query Match: 17.41% Indels: 29
DB: 14 Gaps: 4

US-09-684-215A-4 (1-132) x CB925602 (1-590)

```

```

QY 15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla-----Gly 28
Db 6 GGCTTTGCTATCCCATCATCACTGTACTTAAATCGCTCCTCAGTTAATTGAGTTCCGA 65

QY 29 GlnIleArgSerGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGly 48
Db 66 AAGTTCGTGCGTGGCTTGAATGTGGACTTTCGTCGGGATCCCAATTCATAT----- 119

QY 49 LeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAla 68
Db 120 -----CACGTTAATGTTCCGACGAGGAGCTCTTATCTTAAGTACCTGGGGCAGT 170

QY 69 ProAlaIleSerLeuGlyIleSer-----Thr 77
Db 171 GCTGCAGCCCAAGCAGGCTTCTGCTCCACCGCGCAGGGGTTTGTCTGTTAATATTGTTCTG 230

QY 78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
Db 231 GGTGATATCATCGTTGTCAGTGGACGCGCAACCTGTTAAGGGCAAAATCTGACCTGCTGAGG 290

QY 98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
Db 291 GTTCTGGATGACTATGGCTCGGAGATCAGGTGACCTTGACA----- 332

QY 118 GlyThrArgThrGlyAsnValThrLeu 126
Db 333 ---ATCCGCGCAGGCTCAGAAACCCCTT 356

```

## RESULT 6

```

LOCUS
  A2934428 603 bp DNA linear GSS 24-APR-2001
  BJ_Ba0002108r B. japonicum BAC library Bradyrhizobium japonicum
  genomic, genomic survey sequence.
ACCESSION
  A2934428
  A2934428.1 GI:13776488
KEYWORDS
  GSS.
SOURCE
  Bradyrhizobium japonicum
  Bradyrhizobium japonicum
  Bradyrhizobiales; Alphaproteobacteria; Rhizobiales;
  Tomkins, J.P., Wood, T.C., Stacey, M.G., Loh, J.T., Judd, A., Golcochea
  J.L., Stacey, G., Sadowsky, M.J. and Wing, R.A.
  A marker-denge, sequence-ready map of the Bradyrhizobium japonicum
  genome
  Genome Res. 11 (8), 1434-1440 (2001)
MEDLINE
  21376150
PUBMED
  11483585
COMMENT
  Contact: Wing RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Tel: 864 656 7288
  Fax: 864 656 4293
  Email: rwing@clemson.edu
  Class: BAC ends
  High quality sequence stop: 553.
  Location/Qualifiers
    1..603
    /organism="Bradyrhizobium japonicum"
    /mol_type="genomic DNA"
    /strain="USD110"
    /db_xref="taxon:375"
    /lab_host="E. coli"
    /clone_lib="B. japonicum BAC library"
    /note="Vector: pIndigo536; Site_1: HindIII"

```

## FEATURES

```

source
  1..603
  /organism="Bradyrhizobium japonicum"
  /mol_type="genomic DNA"
  /strain="USD110"
  /db_xref="taxon:375"
  /lab_host="E. coli"
  /clone_lib="B. japonicum BAC library"
  /note="Vector: pIndigo536; Site_1: HindIII"
BASE COUNT 109 a 201 c 203 g 90 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.116 Length: 603
Score: 114.00 Matches: 36

```

Percent Similarity: 43.44%    Conservative: 17  
 Best Local Similarity: 29.51%    Mismatches: 57  
 Query Match: 16.89%    Indels: 12  
 DB: 28    Gaps: 4

US-09-684-215A-4 (1-132) x AZ934428 (1-603)

Oy 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32  
 Db 245 GGCATCGCGTTCGATCCCGCCCAACACCGTGAAGAGCGTGTGGCCACCTCAAGGAC 304  
 Oy 33 GlyGly-----GlySerProThrValHisIleGlyProThrAlaPheLeuGly 48  
 Db 305 AAGGTTGGTTCAGCCCGCGTGGATCGCGTGCAGATTCAGCGGTGACG----- 355  
 Oy 49 LeuGlyValValAspAsnGly-----AsnGlyAlaArgValGlnArgVal 64  
 Db 356 TCGATATCGCGCAGACCGTCCGATGAGAGCCGAGGCGCGTGGTGGCGGAGCCG 415  
 Oy 65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84  
 Db 416 CAGCGAACGGTCCGCGCGCAAGCGCGCATCGAGTCCGCGAGTGCATCACTCGGTC 475  
 Oy 85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 104  
 Db 476 AACGGCAATCCGTCAGAGGACCGCGCGAGTCCCGCCACCATCGCGCGCATCGCGCC 535  
 Oy 105 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 124  
 Db 536 CGTGGATCGTGAGCTTAACGTCTGTCACAAAG---GCCAGGACAATGTCTGTAACCTC 592  
 Oy 125 ThrLeu 126  
 Db 593 ACCCTC 598

# RESULT 7

BQ514888/c BQ514888 758 bp mRNA linear EST 07-MAR-2003  
 DEFINITION EST622303 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMIO81 3' end, mRNA sequence.

ACCESSION BQ514888  
 VERSION BQ514888.1 GI:21373757  
 KEYWORDS EST.

SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.  
 AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.

TITLE Generation of a set of potato cDNA clones for microarray analyses  
 JOURNAL Unpublished  
 COMMENT Other\_ESTs: EST622302

CONTACT: Robin Buell

INSTITUTE: The Institute for Genomic Research

ADDRESS: 9712 Medical Center Dr., Rockville, MD 20850, USA

EMAIL: potato-array@tigr.org

NOTE: This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/

Seq primer: T7.

## FEATURES

source

Location/Qualifiers

1..758

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec or Binjite"

/db\_xref="taxon:4113"

/clone="STMIO81"

/tissue\_type="mixed tissues"

/lab\_host="SOLR"

/clone\_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes tubers, or roots."

BASE COUNT 215 a 181 c 128 g 234 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.19    Length: 758  
 Score: 113.00    Matches: 34  
 Percent Similarity: 42.98%    Conservative: 15  
 Best Local Similarity: 29.82%    Mismatches: 47  
 Query Match: 16.74%    Indels: 18  
 DB: 13    Gaps: 3

US-09-684-215A-4 (1-132) x BQ514888 (1-758)

Oy 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29  
 Db 592 ACATCAGCAGGTGTGGATTGCAATCCCTCTTCAACTGTGTGAGAGTTGTGCCCCAG 533  
 Oy 30 IleArgSerGlyGly-----GlySerProThrValHisIleGlyProThrAla 45  
 Db 532 TTGATCCAATCTGCAAAAGTTCTTCGTCTGCTGGTTGAATATTGAATTCGCTCCAGACCTG 473  
 Oy 46 PheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgValVal 65  
 Db 472 -----ATTGCAACCACTTAATGTTTCAAAATGAGCAGCTGGTTTCTGCTGTAACCT 422  
 Oy 66 GlySerAlaProAlaAlaSerLeuGly----- 74  
 Db 421 GGAATAGTCTCGCAGCAAGCGGAGCTTCTCTACTACCAGGGGTTTTCAGGAAAT 362  
 Oy 75 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 94  
 Db 361 ATAGTGTGTCGATATATTATTGAAGCAGTGTGATGACAAACCTGTTAGGAGTAAAGCAGAG 302  
 Oy 95 MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 108  
 Db 301 TTGTATAAGCCCTGGTAACATAACATAGGTGAAGTT 260

## RESULT 8

BZ561390/c BZ561390

DEFINITION pacs2-164\_3239.y3 pacs2-164 Pseudomonas aeruginosa genomic clone  
 pacs2-164\_3239, genomic survey sequence.

ACCESSION BZ561390

VERSION BZ561390.1 GI:27181349

KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa

ORGANISM Pseudomonas aeruginosa

REFERENCE 1 (bases 1 to 1033)  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol., (2002) In press

COMMENT Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun

Location/Qualifiers

1..1033

/organism="Pseudomonas aeruginosa"

BASE COUNT		188 a	313 c	292 g	236 t	4' others
ORIGIN						
<pre> /mol_type="genomic DNA" /strain="2-164" /db_xref="taxon:287" /clone="pacs2-164_3239" /clone_lib="pacs2-164" /note="clinical isolate 2-164 Whole genomic shotgun library." </pre>						
Alignment Scores:						
Pred. No.:	1,2	Length:	1033			
Score:	106.50	Matches:	41			
Percent Similarity:	44.20%	Conservative:	20			
Best Local Similarity:	29.71%	Mismatches:	56			
Query Match:	15.78%	Indels:	21			
DB:	29	Gaps:	4			
US-09-684-215A-4 (1-132) x BZ561390 (1-1033)						
QY	4	SerAspAsnPheGlnLeuSerGlnGlyGlyGln-----GlyPheAlaIleProIleGly	21			
Db	713	TCCACGATTTTCCACCAGCTCCGCGCGCTCAATGSCCTGTCTTTCGCATTCGCATCGAT	654			
QY	22	GlnAlaMetAlaIle-AlaGlyGlnIleArg-----Se	32			
Db	653	GTCGGCGTACCGTTTCGCGCAGCACGATTGAAGAAAGCCGCCAAGGTCAGTCGGCGTGGC	594			
QY	32	rGlyGlyGlySerProThrValHisIleGlyProThrAlaPheIeu-GlyLeuGlyVal	52			
Db	593	TGGGCGTGGTGATCCAGGAATGAACAAGGATTCGCCGAGTCCTTCGGCCTC-----	541			
QY	52	AlaAspAsnGlyAsnGlyAlaArgValGlnArgValGlySerAlaProAlaAlas	72			
Db	540	-----GACAGCCGTCGCGCGCGTGTGGCGCACTGGTGGAGACGCCCGCGCGCCA	486			
QY	72	erLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSera	92			
Db	485	AGGGTGGCGTGCAGGTGGCGCATGTGATCTCTCAGCCTGAACGCCAGTCAGTCAACGAGT	426			
QY	92	IaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal----	110			
Db	425	CCGCGCAGCTCGCGCAGCTGTGTGGCAACATGAAGCCGCGCGCAAGATCAACCTGGACG	366			
QY	111	-----ThrTrpGlnThrIlySerGlyGlyThrArgThrGlyAsnVal	124			
Db	365	TGATTCACGACGGCCAGCGCAGTCCCTGACCATGGCGGTAGCAGCCTT	316			
RESULT 9	AZ933900	719 bp	DNA	linear	GSS	24-APR-2000
LOCUS	Bj_Ba001B03r B. japonicum BAC library	Bradyrhizobium japonicum				
DEFINITION	genomic, genomic survey sequence.					
ACCESSION	AZ933900.1	GI:13775960				
VERSION	AZ933900					
KEYWORDS	GSS.					
SOURCE	Bradyrhizobium japonicum					
ORGANISM	Bradyrhizobium japonicum					
REFERENCE	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;					
AUTHORS	Bradyrhizobiaceae; Bradyrhizobium.					
TITLE	1 (bases 1 to 719)					
JOURNAL	Tomkins, J.P., Wood, T.C., Stacey, M.G., Loh, J.T., Judd, A., Goicoechea,					
MEDLINE	J.L., Stacey, G., Sadowsky, M.J. and Wing, R.A.					
PUBLISHED	A marker-dense, sequence-ready map of the Bradyrhizobium japonicum					
COMMENT	Genome Res. 11 (8), 1434-1440 (2001)					
	21376150					
	11483585					
	Contact: Wing RA					
	Clemson University Genomics Institute					
	Clemson University					
	100 Jordan Hall, Clemson, SC 29634, USA					
	Tel: 864 656 7288					

```

Fax: 864 656 4293
Email: rwing@clemons.edu
Class: BAC ends
High quality sequence stop: 671.
FEATURES             source
    Location/Qualifiers
        1..719
            /organism="Bradyrhizobium japonicum"
            /mol_type="genomic DNA"
            /strain="USDAL10"
            /db_xref="taxon:375"
            /lab_host="E. coli"
            /clone_lib="B. japonicum BAC library"
            /note="Vector: pIndigo36; Site_1: HindIII"
BASE COUNT      127 a      245 c      235 g      111 t
ORIGIN
Alignment Scores:
Pred. No.:      1.55      Length:      719
Score:          103.50    Matches:      34
Percent Similarity: 37.86%    Conservative: 19
Best Local Similarity: 24.29%    Mismatches: 62
Query Match:      15.33%    Indels:      25
DB:               26      Gaps:      4

US-09-684-215A-4 (1-132) x AZ933900 (1-719)
QY      13 GlyGInGlyPheAlaIleProIleGlyClnAlaMetAlaIleAlaGlyGlnIleArgSer 32
Db      44 GGCATCGGCTTTGGATCTCTTGTCACATGAGCGCGGTGCTGCTCGCTCCGCCACAGGCG 103
QY      33 GlyGly-----GlySerProThrValHisIleGlyProThr 44
Db      104 GCGCGCAAGCGGTGAAGCGTCCCTGGCTCGGCGCGAAGTTGACGGCGGTGACGCCCGAG 163
QY      45 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 64
Db      164 ATCGCGCGAAAGCCCTCGGTTG-----CGTTCGCGCACCGCGCGCTGTCGCGAGCGTG 217
QY      65 ValGlySerAlaProAlaAserLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
Db      218 GTCTCGAAGCGTCCGCGACGACGAGCGCGGCTGAAATCTCCGATCTGATCACCGGGGATC 277
QY      85 AspGlyAlaProIleAsn-----GlySerProThrValHisIleGlyProThr 90
Db      278 GAGCGCCAGACCGTGGATGATCCCAATGCCTTCGACTACCGTTCGCGAGCGGTCCACTC 337
QY      91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
Db      338 GCGCGTTCGCGCGAGATCGACGGGCGACGCGGCGGCGAAGCGCGTCAAG---CTCGCGATC 394
QY      111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 130
Db      395 GCGTGGAGACCGCACCGGCGCGCGCGCAACGAGCTCGTCACTACGCGGCTTCGCCG 454

RESULT 10
LOCUS   B2549047/c
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```



100

..... 11



```

ACCESSION   AI597611
VERSION     AI597611.1  GI:4606659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 423)
AUTHORS    NCi/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGP), Tumor Gene Index
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapsb-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.lnl.gov/bbrp/image/image.html
            Seq primer: 40UP from Gibco.
FEATURES    source
            Location/Qualifiers
                1..423
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:2167707"
                /tissue_type="anaplastic oligodendroglioma"
                /lab_host="DH10B"
                /clone_lib="NCI_CGAP_Brn25"
                /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
                modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                strand cDNA was primed with a Not I - oligo(dT) primer [5'
                TGTTACCAATGTAAGTGGAGCGCGCATAGGTTTTTTTTTTTTTTTTTTT
                T 3']; double-stranded cDNA was ligated to Eco RI
                adaptors (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of the modified pT73 vector.
                Library is normalized, and was constructed by Bento
                Soares and M. Fatima Bonaldo."
BASE COUNT  65 a 111 c 139 g 108 t
ORIGIN
Alignment Scores:
Pred. No.: 2.31 Length: 423
Score: 99.00 Matches: 35
Percent Similarity: 44.35% Conservative: 20
Best Local Similarity: 28.23% Mismatches: 43
Query Match: 14.67% Indels: 26
DB: 9 Gaps: 7
US-09-684-215A-4 (1-132) x AI597611 (1-423)
QY 7 PheGlnLeuSerGln-----GlyGlnGlnGlyPheAlaIleProIleGlyGlnAla 23
    |||||
Db 39 TTTGAAAAGGCCAAAGGGCCCTGGGGCAACAGGACAGGCGCGGTTCTCAGGGG 98
QY 24 MetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThr---ValHisIleGly 42
    |||||
Db 99 TCAGGGGCATTGGGCAGATCGCTGTAGTGGGGGGCCACCTCCCAAGTCCTTTGTGTCA 158
    |||||
QY 43 ProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArgValGln 62
    |||||
Db 159 CCCTGGGTGCTCAGGCACAGGT-----GGCACTGGCTCAAAACAGG 203
QY 63 ArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThr 82
    |||||
Db 204 TAAGCTTCGGGGCGCTACCTTCACGAGTCTTGA-----ACAGGCTCCTCCTGTGTC 257
QY 83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102

```

258 AGCATTTGGGGTTCGTCGGTCTGCCACGAGTTATG----- 293  
 103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGly-----Gly 118  
 294 CCCCCGGA-----TCTCTTTTGGCCACCCTGGGGGCTGAGCCTTCTCAGGG 344  
 119 ThrArgThrGly 122  
 345 ACCCTTCTGGG 356  
 RESULT 15  
 BI350520/c  
 LOCUS  
 DEFINITION fr32b05.y1 zebrafish adult brain Danio rerio cDNA clone  
 IMAGE:4955073 5' similar to SW:ACT2\_FUGRU P53485 ACTIN, CYTOPLASMIC  
 2 : mRNA sequence.  
 ACCESSION BI350520 499 bp mRNA linear EST 26-JUL-2002  
 VERSION fr32b05.y1 zebrafish adult brain Danio rerio cDNA clone  
 KEYWORDS IMAGE:4955073 5' similar to SW:ACT2\_FUGRU P53485 ACTIN, CYTOPLASMIC  
 SOURCE 2 : mRNA sequence.  
 ORGANISM BI350520.1 GI:15044966  
 EST.  
 Danio rerio (zebrafish)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
 ; Cyprinidae; Danio.  
 REFERENCE 1 (bases 1 to 499)  
 AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
 S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood  
 K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
 Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,  
 Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.  
 and Wilson,R.  
 TITLE WashU zebrafish EST Project 1998  
 JOURNAL Unpublished  
 COMMENT Contact: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbrafish@watson.wustl.edu  
 cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:  
 Matthew Clark. DNA Sequencing by: Washington University Genome  
 Sequencing Center Clone distribution: Genome Systems, St. Louis,  
 Missouri (web address: www.genomesystems.com) (email contact:  
 info@genomesystems.com) and Research Genetics, Huntsville, Alabama  
 (web address: www.resgen.com) (email contact: info@resgen.com) and  
 RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address:  
 www.rzpd.de)  
 Seq primer: 17  
 High quality sequence stop: 301.  
 Location/Qualifiers  
 1..499  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"  
 /clone="IMAGE:4955073"  
 /sex="mixed male and female"  
 /tissue\_type="brain"  
 /dev\_stage="adult"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="zebrafish adult brain"  
 /note="vector: pZIPLOX; Site\_1: NotI; Site\_2: SalI;  
 Original library was constructed in lambdaZIPLOX. Mass  
 excision of the cDNA library was performed to yield  
 pZIPLOX plasmids. Insert check was done in original  
 library."  
 BASE COUNT 101 a 156 c 125 g 117 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.95 Length: 499  
 Score: 97.50 Matches: 38  
 Percent Similarity: 39.29% Conservative: 17

Best Local Similarity: 27.14% Mismatches: 52  
Query Match: 14.44% Indels: 33  
DB: 12 Gaps: 6

US-09-684-215A-4 (1-132) x BI350520 (1-499)

QY	9	LeuSerGlnGly-----GlyGlnGlyPheAlaIleProIle	20
DB	430	CTGGAACAGGGCTCTGGACACCTGAACCTTTTCATTGCCAATGGTGATGACCTGTCGGTC	371
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40
DB	370	AGGCAGCTGGTAGCTCTT-----CTCCAGGGAGGAGGAGGAGCAGCGGTGCCCAT	320
QY	41	Ile-----GlyProThrAlaPheLeuGlyValValAspAsnAsn	55
DB	319	CTCCTGCTCAAGTCAAGGCCACATAGCAGAGCTTCTCTTGATGTCACGGACAAT---	263
QY	56	GlyAsnGlyAlaArgValGlnArgValGlySerAlaProAlaAlaSerLeuGlyIle	75
DB	262	-----TTCCCTCTCAGCTGTGGTGAAGCTGTAGCCTCTCTCGGTGAG	218
QY	76	SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet	95
DB	217	GATCTTCATGAGGTAGTCAGTCAGGTCACGGCCAGCCAAAGTCCACAGGAGGATGGCARG	158
QY	96	AlaAspAla-----LeuAsnGlyHisHisProGlyAspValIleSerValThrTrp	112
DB	157	GGCAGGGCGTAAACCCTCGTAGATGGGCACAGTGTGGTGACACCATCACCATTAGTCCAT	98
QY	113	GlnThrLysSerGlyGlyThrArg-----ThrGlyAsnVal	124
DB	97	CACGATACCATGGTGTACGACAGGAGGATACAGGAACACAGCAGCGCTGGATGGCAACGTG	38

Search completed: September 5, 2003, 11:54:34  
Job time : 2287.41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:24:34 ; Search time 58.2414 Seconds  
(without alignments)  
1390.474 Million cell updates/sec

Title: US-09-684-215A-17

Perfect score: 148  
Sequence: 1 TAASDNFQLSQGQGFPAIPGQAWAIAGQI 30

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xip  
-Q/cgn2.1/USPTO.spool\_P/US09684215/runat\_05092003\_072201\_548/app\_query.fasta\_1.853  
-DB=N.Geneseq.19Jun03 -CPMT=fastap -SUFFIX=ring -MINMATCH=0 -1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blos62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -RCGN\_1\_1\_626 -runat\_05092003\_072201\_548 -NCFU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.19Jun03.\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148	100.0	186	24	ABK39777 DNA encoding Ral2S
2	148	100.0	186	25	ACAL12106 Human His-tagged R
3	148	100.0	186	25	ACA03292 Lung cancer therap
4	148	100.0	396	23	AA40769 Nucleotide sequenc
5	148	100.0	399	24	AB571801 Human DNA encoding
6	148	100.0	447	18	AA791486 Mycobacterium tube
7	148	100.0	447	18	AA791403 Mycobacterium tube
8	148	100.0	447	19	AAV64450 M. tuberculosis im
9	148	100.0	447	19	AAV44342 Mycobacterium tube
10	148	100.0	447	20	AA219252 M. tuberculosis an
11	148	100.0	447	20	AA219040 M. tuberculosis re
12	148	100.0	447	22	AA03780 M. tuberculosis DN
13	148	100.0	447	24	AA047080 Mycobacterium tube
14	148	100.0	447	24	AA028339 Mycobacterium spec
15	148	100.0	672	23	AA040772 Nucleotide sequenc
16	148	100.0	675	22	AA564132 Human /M. tubercul
17	148	100.0	675	22	AAH93896 Ral2-P510S-C const
18	148	100.0	675	24	ABL95503 Ral2- P510S-C const
19	148	100.0	675	25	ACA59940 Prostate cancer th
20	148	100.0	702	20	AA220206 Mycobacterium tube
21	148	100.0	702	23	AA040770 Nucleotide sequenc
22	148	100.0	702	24	ABK14140 DNA encoding antig
23	148	100.0	822	24	ABK39769 DNA encoding lung
24	148	100.0	822	25	ACAL2098 Human lung cancer
25	148	100.0	822	25	ACA03284 Lung cancer therap
26	148	100.0	861	24	ABK39776 DNA encoding Ral2S
27	148	100.0	861	25	ACAL12105 Human Ral2S-L985P
28	148	100.0	861	25	ACA03291 Lung cancer therap
29	148	100.0	894	24	ABK27798 Human colon cancer
30	148	100.0	900	24	ABQ92443 Human lung cancer
31	148	100.0	900	24	ABQ92437 Ral2/C-terminal po
32	148	100.0	915	22	AA564141 Human /M. tubercul
33	148	100.0	915	22	AAH93905 Ral2-P775P-ORF3 co
34	148	100.0	915	24	ABL95512 Ral2- P775P-ORF3 c
35	148	100.0	915	25	ACA59949 Prostate cancer th
36	148	100.0	945	24	ABK39768 DNA encoding lung
37	148	100.0	945	25	ACAL2097 Human lung cancer
38	148	100.0	945	25	ACA03283 Lung cancer therap
39	148	100.0	1002	24	AA047077 Mycobacterium tube
40	148	100.0	1002	24	AA047078 Mycobacterium tube
41	148	100.0	1002	24	AA028336 Mycobacterium sp.
42	148	100.0	1002	24	AA028337 Mycobacterium spec
43	148	100.0	1012	24	ABQ92442 Human lung cancer
44	148	100.0	1012	24	ABL49256 Ral2/N-terminal po
45	148	100.0	1035	24	ABK69714 DNA encoding human

ALIGNMENTS

RESULT 1

ABK39777  
ID ABK39777 standard; cDNA; 186 Bp.

AC ABK39777;

DT 21-MAY-2002 (first entry)

XX DNA encoding Ral2S-L985Pex peptide.

XX Lung tumour; cancer; T cell; immune response stimulator;

XX cytostatic; gene; ss.

XX Homo sapiens.

XX Synthetic.

XX WO200204514-A2.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

XX 17-JAN-2002.  
 XX 10-JUL-2001; 2001WO-US22058.  
 XX 11-JUL-2000; 2000US-0614124.  
 PR 29-AUG-2000; 2000US-0651563.  
 PR 08-SEP-2000; 2000US-0658824.  
 PR 26-SEP-2000; 2000US-0671325.  
 PR 06-OCT-2000; 2000US-0677419.  
 PR 30-OCT-2000; 2000US-0702705.  
 PR 13-DEC-2000; 2000US-0736457.  
 PR 03-MAY-2001; 2001US-0849626.  
 XX (CORI-) CORIXA CORP.  
 XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;  
 PI Marnerakis M, Carter D, Fanger GR, Vedvick IS, Bangur CS;  
 PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
 XX WPI; 2002-164634/21.  
 DR P-PSDB; AAU85595.  
 XX Novel polynucleotide encoding a lung tumour polypeptide useful for  
 PT stimulating and/or expanding T cells specific for a tumour protein  
 XX  
 PS Claim 1; SEQ ID No 1879; 223pp; English.  
 CC The invention describes an isolated polynucleotide and polypeptide  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein for determining the presence of a cancer in a patient. A  
 CC composition containing the polynucleotide and/or polypeptide is useful  
 CC for treating a lung cancer in a patient. The polypeptide is useful for  
 CC removing tumour cells from a biological sample. The polynucleotide is  
 CC also useful as probe or primer to detect the level of mRNA encoding a  
 CC tumour protein. This sequence encodes a lung tumour associated protein  
 CC or protein fragment, described in the method of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,99e-15 Length: 186  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-684-215A-17 (1-30) x ABK39777 (1-186)  
 QY 1 ThrAlaAlaSerAspAspPheGlnLeuSerGlnGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 22 AGCGCGCGTCCGATPAACGTCCAGCTGCCAGGTGGCAGGATCCCATTCGCATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 DB 82 GGCAGCGCGATGGCGATCGCGGCCAGATC 111  
 RESULT 2  
 ID ACA12106  
 AC ACA12106 standard; cDNA; 186 BP.  
 XX ACA12106;  
 AC ACA12106;  
 XX 06-JUN-2003 (first entry)  
 DT Human His-tagged Ral2S-L985pEx fusion protein cDNA.  
 DE Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
 KW

KW T cell expansion; CD4; CD8; RAL2; gene.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX US2002197669-A1.  
 PN 26-DEC-2002.  
 PD 03-MAY-2001; 2001US-0849626.  
 PF 13-DEC-2000; 2000US-0736457.  
 PR (BANG/) BANGUR C S.  
 PA (FANG/) FANGER G R.  
 PA (WANG/) WANG A.  
 PA (WANG/) WANG T.  
 PA (SWIT/) SWITZER A P.  
 PA (MCNE/) MCNEILL P D.  
 PA (CLAP/) CLAPPER J D.  
 XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
 PI Clapper JD;  
 PI WPI; 2003-352750/33.  
 DR P-PSDB; ABU69570.  
 XX Novel lung cancer polynucleotide encoding lung cancer protein, useful  
 PT for detecting the presence of lung cancer in a patient, and in  
 PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer  
 XX  
 PS Example 10; Page -: 72pp; English.  
 CC The invention relates to a polynucleotide encoding a lung tumour protein,  
 CC comprising a sequence selected from any of the 14 sequences  
 CC mentioned in the specification, or a sequence (S2) mentioned in  
 CC contiguous residues of S1, sequences consisting of at least 20  
 CC having 75%, preferably 90%, identity to S1, or degenerate variants of  
 CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)  
 CC selected from any one of the 4 amino acid sequences mentioned in the  
 CC specification, a sequence encoded by the polynucleotide, or sequences  
 CC having at least 70%, preferably 90%, identity to a sequence encoded by  
 CC the polynucleotide), an expression vector comprising the polynucleotide  
 CC operably linked to an expression control sequence, a host cell  
 CC transformed or transfected with the vector, an isolated antibody (or its  
 CC antigen-binding fragment) that specifically binds to the polypeptide,  
 CC detecting the presence of a cancer in a patient, a fusion protein  
 CC comprising the polypeptide, an oligonucleotide that hybridises to  
 CC S1 under moderately stringent conditions, stimulating and/or expanding T  
 CC cells specific for a tumour protein (comprising contacting T cells with  
 CC the polynucleotide, protein or antigen-presenting cells, under conditions  
 CC and for a time sufficient to permit the stimulation and/or expansion of T  
 CC cells) and inhibiting the development of a cancer in a patient (by  
 CC incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with the  
 CC polynucleotide, protein or antigen presenting cells that express the  
 CC polynucleotide, such that T cells proliferate, administering to the  
 CC patient an effective amount of the proliferated T cells, and thus  
 CC inhibiting the development of a cancer in the patient. The  
 CC polynucleotide, protein and cells are useful in a composition for  
 CC stimulating an immune response in a patient, and for treating a cancer in  
 CC a patient (particularly lung cancer). The oligonucleotide is useful for  
 CC determining the presence of a cancer in a patient. The protein and  
 CC oligonucleotides are useful in pharmaceutical compositions, e.g.  
 CC vaccines. The polynucleotide is also useful as a probe or primer for  
 CC nucleic acid hybridisation, and in the design and preparation of  
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
 CC proteins in tumour cells. An amplified portion of the polynucleotide is  
 CC useful for isolating a full-length gene from a suitable library.  
 CC The present sequence encodes a fusion protein of human RAL2 with  
 CC the protein product of a cDNA (full length, extended or partial)  
 CC isolated from a library derived from lung tumour/cancer cells.  
 CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic  
CC format directly from the USPTO  
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.

XX Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 other;

Alignment Scores: 2.99e-15 Length: 186  
Pred. No.: 148.00 Matches: 30  
Score: 148.00  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-684-215A-17 (1-30) x ACA12106 (1-186)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProile 20  
Db 22 ACGCCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
Db 82 GGGCAGCGGATGGCGATCGCGGGCCAGATC 111

## RESULT 3

ACA03292  
ID ACA03292 standard; cDNA: 186 BP.

XX ACA03292;

XX 22-MAY-2003 (first entry)

XX Lung cancer therapy and diagnosis associated cDNA #1767.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer;

XX gene; ss.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-0902941.

XX 30-JUN-1999; 99US-0346492.

XX 15-OCT-1999; 99US-0419356.

XX 17-DEC-1999; 99US-0466867.

XX 30-DEC-1999; 99US-0476300.

XX 06-MAR-2000; 2000US-0519642.

XX 22-MAR-2000; 2000US-0533077.

XX 10-APR-2000; 2000US-0546259.

XX 27-APR-2000; 2000US-0560406.

XX 05-JUN-2000; 2000US-0589184.

XX 11-JUL-2000; 2000US-0614124.

XX 29-AUG-2000; 2000US-0651563.

XX 08-SEP-2000; 2000US-0658824.

XX 26-SEP-2000; 2000US-0671325.

XX 06-OCT-2000; 2000US-0677419.

XX 30-OCT-2000; 2000US-0702705.

XX 13-DEC-2000; 2000US-0736457.

XX 03-MAY-2001; 2001US-0849626.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;

XX Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.

XX New polynucleotide, useful for preparing a composition for treating or

XX inhibiting development of cancer, e.g. lung cancer

XX

PS Example 10; SEQID NO 1879; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32  
CC sequences, complement or degenerate variants of them. The polynucleotide  
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
CC for treating or inhibiting development of cancer, e.g. lung cancer.  
CC This sequence represents a polynucleotide associated with the  
CC compositions and methods for the therapy and diagnosis of lung cancer.

XX Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 other;

Alignment Scores: 2.99e-15 Length: 186  
Pred. No.: 148.00 Matches: 30  
Score: 148.00  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-684-215A-17 (1-30) x ACA03292 (1-186)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProile 20  
Db 22 ACGCCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCCGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

Db 82 GGGCAGCGGATGGCGATCGCGGGCCAGATC 111

## RESULT 4

AAL40769  
ID AAL40769 standard; DNA; 396 BP.

XX AAL40769;

XX 03-OCT-2002 (first entry)

XX Nucleotide sequence encoding Ral2 protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
XX vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
XX immunogen; cytokine; gene; ds.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 1..396

XX /\*tag= a

XX /product= "Ral2 protein"

XX /note= "No start or stop codon"

XX WO200125401-A2.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27652.

XX 07-OCT-1999; 99US-0158585.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Guderian J;

XX WPI; 2001-266299/27.

XX P-PSDB; RAO22138.

XX Recombinant nucleic acid molecule for producing high yield expression

XX of desired fusion polypeptides, encodes fusion polypeptide comprising

XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide

XX Claim 1; Fig 2; 39pp; English.

XX

CC The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
CC kDa C-terminal fragment of serine protease antigen MTB32A of  
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
CC The recombinant fusion nucleic acids and polypeptides are useful for  
CC providing stable and high yield expression of fusion polypeptides of both  
CC eukaryotic and prokaryotic origin and to encode a protein product for use  
CC as an antigen for detecting serum antibodies. The presence of serum  
CC antibodies to M. tuberculosis antigens in an individual indicates that  
CC the individual is infected with it. The fusion polypeptides are useful as  
CC sources of proteins for monitoring binding of serum antibodies to fusion  
CC proteins and as an immunogen to induce and/or enhance immune responses.  
CC The coding sequences can be ligated with a coding sequence of another  
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
CC represents the DNA encoding the Ral2 protein.

XX Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 other;

#### Alignment Scores:

Pred. No.:	7.59e-15	Length:	396
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	23	Indels:	0
DB:		Gaps:	0

US-09-684-215A-17 (1-30) x AAL40769 (1-396)

QY	1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB	1 ACGGCGCGTCCGATAAATCCAGCTGCCAGGTGGCAGGGATTGCCATTCGGATC 60
QY	21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB	61 GGGCAGGCGATGGCGATCGCGGCCAGATC 90

#### RESULT 5

ID	ABS71801	standard; DNA; 399 BP.
XX	ABS71801;	
XX	02-DEC-2002 (first entry)	
XX	Human DNA encoding a mammaglobin/Ral2 fusion protein.	
XX	Human; epitope; mammaglobin; breast cancer; cytostatic; T cell; CD4+;	
KW	CD8+; antigen; Ral2; ds; gene.	
XX	Chimeric - Homo sapiens.	
OS	Chimeric - Mycobacterium tuberculosis.	
XX	WO200253017-A2.	
XX	11-JUL-2002.	
XX	08-JAN-2002; 2002WO-US03057.	
PR	08-JAN-2001; 2001US-0757417.	
PR	08-NOV-2001; 2001US-0008045.	
XX	(CORI-) CORIXA CORP.	
XX	Fling SP, Foy TM, Clapper JD, Wang A, Johnson JC, McNeill PD;	
PI	Sutherland RA;	
XX	WPI; 2002-706844/76.	
DR	P-PSDB; ABG94684.	
XX	Novel polypeptides comprise one or more human mammaglobin epitopes and	
PT	polynucleotides encoding the polypeptides, useful for preventing and	
PT	treating breast cancers -	

XX Claim 20; Page 115; 121pp; English.  
PS The invention relates to an isolated polypeptide (I) comprising 7-30  
XX consecutive amino acid residues of human mammaglobin, where one or  
CC more mammaglobin-specific T cells specifically reacts with (I).  
CC Also included are (1) a composition comprising (I), in combination with  
CC a physiologically acceptable carrier or immunostimulant; (2) a diagnostic  
CC kit, comprising (I) and a detection reagent comprising a reporter group;  
CC (3) removing tumour cells from a biological sample, by contacting a  
CC biological sample with T cells that specifically react with (I), under  
CC conditions and for a time sufficient to permit the removal of cells  
CC expressing mammaglobin or a peptide epitope from the sample; (4) an  
CC isolated T cell population (II), comprising T cells prepared using (I);  
CC and (5) a composition (C2) comprising a polynucleotide comprising a  
CC sequence encoding human mammaglobin (or variant or tagged with an  
CC affinity tag), or a polypeptide comprising a human mammaglobin (or  
CC variant or tagged with an affinity tag), in combination with an  
CC immunostimulant. The peptide is useful for inhibiting the development of  
CC breast cancer in a patient, by incubating CD4+ and/or CD8+ T cells  
CC isolated from a patient with a mammaglobin epitope, such that T cells  
CC proliferate, administering the proliferated T cells to the patient,  
CC optionally cloning at least one proliferated T cell and administering at  
CC least one cloned cell, and thus inhibiting the development of breast  
CC cancer in the patient. The compositions and methods are useful for  
CC inhibiting the development of breast cancer in a patient. The  
CC present sequence encodes a human mammaglobin/Ral2 fusion protein.  
XX Sequence 399 BP; 109 A; 96 C; 92 G; 102 T; 0 other;

#### Alignment Scores:

Pred. No.:	7.66e-15	Length:	399
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	24	Indels:	0
DB:		Gaps:	0

US-09-684-215A-17 (1-30) x ABS71801 (1-399)

QY	1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB	22 ACGGCGCGTCCGATAAATCCAGCTGCCAGGTGGCAGGGATTGCCATTCGGATC 81
QY	21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB	82 GGGCAGGCGATGGCGATCGCGGCCAGATC 111

#### RESULT 6

AAAT91466	standard; DNA; 447 BP.
ID	AAAT91466
XX	AAAT91466;
XX	07-JAN-1998 (first entry)
XX	Mycobacterium tuberculosis antigen TBRa12 encoding DNA.
XX	Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW	skin testing; M. tuberculosis; ss.
XX	Mycobacterium tuberculosis.
XX	Key
FT	Location/Qualifiers
FT	11..409
FT	/tag= a
FT	/product= Antigen_TBRa12
XX	WO9709428-A2.
XX	13-MAR-1997.
XX	30-AUG-1996; 96WO-US14674.

```

XX PR 12-JUL-1996; 96US-0680574.
XX PR 01-SEP-1995; 95US-0523436.
XX PR 22-SEP-1995; 95US-0533634.
XX PR 22-MAR-1996; 96US-0620874.
XX PR 05-JUN-1996; 96US-0659683.
XX PA (CORI-) CORIXA CORP.
XX PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
XX PI Twardzik DR, Vedvick TH;
XX DR WPI; 1997-192903/17.
XX DR P-PSDB; AAW32422.
XX PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
XX PT useful in vaccines for prevention or treatment of tuberculosis, also
XX PT for diagnosis
XX PS Claim 3; Page 50; 168pp; English.
XX CC A new immunogenic polypeptide has been developed comprising an
XX CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
XX CC its variant differing only in conservative substitutions and/or
XX CC modifications). The present sequence encodes a specifically claimed
XX CC M.tuberculosis antigen, TbrA12. The immunogenic protein, and fusion
XX CC proteins containing one or more of the proteins or one of the proteins
XX CC plus ESAT-6, are useful in vaccines, preferably when formulated with a
XX CC non-specific adjuvant, to induce an immune response against
XX CC M.tuberculosis (for treatment or prevention).
XX SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:
Pred. No.: 8.81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-684-215A-17 (1-30) x AAT91466 (1-447)
Qy 1 ThrAlaLaserAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGTCAGGTTGGCAGGGATTCCGCATTCGCATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGCAGCGGATGCGGATCGCGGCCAGATC 100

RESULT 7
AAT91403
ID AAT91403 standard; DNA; 447 BP.
XX AC AAT91403;
XX DT 12-JAN-1998 (first entry)
XX DE Mycobacterium tuberculosis antigen TbrA12 encoding DNA.
XX KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
XX KW skin testing; M.tuberculosis; ss.
XX OS Mycobacterium tuberculosis.
XX FH Key Location/Qualifiers
XX FT CDS 11..409
XX FT /*tag= a
XX FT /product= Antigen_TbrA12
XX PN W09709429-A2.

```

```

PD 13-MAR-1997.
XX 30-AUG-1996; 96WO-US14675.
XX 12-JUL-1996; 96US-0680573.
XX 01-SEP-1995; 95US-0523435.
XX 22-SEP-1995; 95US-0532136.
XX 22-MAR-1996; 96US-0620280.
XX 05-JUN-1996; 96US-0658800.
XX (CORI-) CORIXA CORP.
XX PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
XX PI Twardzik DR, Vedvick TH;
XX DR WPI; 1997-192904/17.
XX DR P-PSDB; AAW32354.
XX PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens
XX PT - useful for diagnosis of M. tuberculosis infection
XX PS Claim 3; Page 49-50; 190pp; English.
XX CC A new immunogenic polypeptide has been developed comprising an
XX CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
XX CC its variant differing only in conservative substitutions and/or
XX CC modifications). The present sequence encodes a specifically claimed
XX CC M.tuberculosis antigen, TbrA12. The immunogenic polypeptide can be
XX CC used to diagnose M.tuberculosis infection by forming complexes with
XX CC specific antibodies in the sample. Fragments of DNA encoding the
XX CC immunogenic polypeptide can be used as diagnostic primers or probes
XX CC and agents that bind to the antigen, especially monoclonal antibodies
XX CC or equivalent polyclonal antibodies, are also used for diagnosis.
XX SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:
Pred. No.: 8.81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-684-215A-17 (1-30) x AAT91403 (1-447)
Qy 1 ThrAlaLaserAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGTCAGGTTGGCAGGGATTCCGCATTCGCATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGCAGCGGATGCGGATCGCGGCCAGATC 100

RESULT 8
AAV64450
ID AAV64450 standard; DNA; 447 BP.
XX AC AAV64450;
XX DT 27-JAN-1999 (first entry)
XX DE M. tuberculosis immunogenic polypeptide TbrA12 DNA.
XX KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
XX KW vaccine; pharmaceutical; infection; diagnosis; ss.
XX OS Mycobacterium tuberculosis.
XX PN W09816646-A2.
XX PD 23-APR-1998.

```

```
PF 07-OCT-1997; 97WO-US18293.
XX
PR 13-MAR-1997; 97US-0818112.
PR 11-OCT-1996; 96US-0730510.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI: 1998-261042/23.
DR P-PSDB; AAW81857.
XX
XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and for diagnosis, treatment and prevention of tuberculosis
XX
PS Claim 3; Page 62-63; 230pp; English.
XX
CC This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC tuberculosi (MT) antigen which can be used in a method for inducing
CC protective immunity against tuberculosis (TB). This sequence can be
CC formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis.
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;
XX
Alignment Scores:
Pred. No.: 8,81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19 Indels: 0
DB: Gaps: 0
XX
US-09-684-215A-17 (1-30) x AAV64450 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGGCCGCTCGATACACTTCCAGCTGTCCAGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100
XX
RESULT 9
AAV44342
ID AAV44342 standard; DNA; 447 BP.
XX
AC AAV44342;
XX
DT 09-NOV-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen TBRa12 DNA.
XX
KW Tuberculosis; infection; diagnosis; antigen; TBRa12; ss.
XX
OS Mycobacterium tuberculosis strain H37Ra.
XX
FH Location/Qualifiers
FT 11..406
FT /*tag= a
XX
XX WO9816645-A2.
XX
XX 23-APR-1998.
XX
PF 07-OCT-1997; 97WO-US18214.
XX
PR 13-MAR-1997; 97US-0818111.
PR 11-OCT-1996; 96US-0729622.
XX
XX
```

```
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI: 1998-251292/22.
DR P-PSDB; AAW64294.
XX
XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and diagnosis of tuberculosis
XX
PS Claim 3; Page 61; 250pp; English.
XX
CC This DNA sequence codes for Mycobacterium tuberculosis soluble
CC antigen TBRa12 (see AAW64294). It was isolated from a M. tuberculosis
CC strain H37Ra expression library with rabbit anti-sera raised
CC against M. tuberculosis supernatant. The invention relates to
CC compositions and methods for diagnosing tuberculosis. It provides
CC polypeptides (see AAW64291-W64379) comprising an antigenic portion of
CC a soluble M. tuberculosis antigen, or an immunogenic portion of an
CC M. tuberculosis antigen, as well as DNA sequences encoding such
CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic
CC kits for detecting M. tuberculosis infection in a patient using
CC these polypeptides, antibodies or oligonucleotide probes and
CC primers, for the diagnosis of tuberculosis.
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;
XX
Alignment Scores:
Pred. No.: 8,81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19 Indels: 0
DB: Gaps: 0
XX
US-09-684-215A-17 (1-30) x AAV44342 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGGCCGCTCGATACACTTCCAGCTGTCCAGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100
XX
RESULT 10
AAZ19252
ID AA219252 standard; DNA; 447 BP.
XX
AC AA219252;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis antigen TBRa12 DNA sequence.
XX
KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test; ss.
XX
OS Mycobacterium tuberculosis.
XX
XX WO9942076-A2.
XX
XX 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US03268.
XX
PR 05-MAY-1998; 98US-0072967.
PR 18-FEB-1998; 98US-0025197.
XX
XX
```



PA (CORI-) CORIXA CORP.  
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;  
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;  
XX  
XX WPI; 1999-527409/44.  
DR P-PSDB; AAY39096.  
XX  
XX New antigens from Mycobacterium tuberculosis useful in diagnostic  
PT skin tests and protective or therapeutic vaccines or compositions  
PT  
XX  
XX Claim 3; Page 73; 299pp; English.  
XX  
XX The present invention describes polypeptides comprising an immunogenic  
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described  
CC are vaccines and fusion protein containing M. tuberculosis Ag's.  
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and  
CC other polypeptides fragments, can be used in pharmaceutical compositions  
CC or vaccines to generate a protective or therapeutic immune response to  
CC M. tuberculosis and as reagents in skin tests for diagnosis of  
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion  
CC by, T, B or natural killer cells and/or macrophages in  
CC tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to  
CC AAY39225 are used in the exemplification of the present invention.  
XX  
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;  
  
Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
  
US-09-684-215A-17 (1-30) x AAZ19252 (1-447)  
  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
Db 11 ACGCCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCGGATC 70  
  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
Db 71 GGCAGCGCATGCGATCGCGGCCAGATC 100  
  
RESULT 11  
AAZ19040  
ID AAZ19040 standard; DNA: 447 BP.  
XX  
AC AAZ19040;  
XX  
DT 05-NOV-1999 (first entry)  
XX  
DE M. tuberculosis recombinant antigen DNA encoding TbrA12.  
XX  
KW Antigen; diagnosis; detection; infection; antibody; immunisation;  
KW vaccine; immunity; ss.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO9942118-A2.  
XX  
PD 26-AUG-1999.  
XX  
PF 17-FEB-1999; 99WO-US03265.  
XX  
PR 05-MAY-1998; 98US-0072596.  
PR 18-FEB-1998; 98US-0024753.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;  
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

XX  
DR WPI; 1999-527416/44.  
DR P-PSDB; AAY38959.  
XX  
XX New polypeptide comprising antigenic portions of M. tuberculosis  
PT  
XX Claim 1; Page 103; 323pp; English.  
XX  
XX This invention describes novel recombinant antigens and their encoding  
CC nucleic acids derived from Mycobacterium tuberculosis. The novel  
CC polypeptides are useful for detecting M. tuberculosis infection in a  
CC biological sample by detecting antibodies which bind with the  
CC polypeptides, and are useful as vaccines for immunizing against  
CC M. tuberculosis infection. The new detection methods are needed as  
CC current vaccination strategies do not provide 100% immunity.  
XX  
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;  
  
Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
  
US-09-684-215A-17 (1-30) x AAZ19040 (1-447)  
  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
Db 11 ACGCCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCGGATC 70  
  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
Db 71 GGCAGCGCATGCGATCGCGGCCAGATC 100  
  
RESULT 12  
AAZ03780  
ID AAZ03780 standard; DNA: 447 BP.  
XX  
AC AAZ03780;  
XX  
DT 29-AUG-2001 (first entry)  
XX  
DE M. tuberculosis DNA encoding a partial antigen TbrA12.  
XX  
KW TbrA12; antigen; vaccine; tuberculosis; AIDS;  
KW acquired immunodeficiency disease; ss.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN Key Location/Qualifiers  
FT CDS 11..409  
FT /\*tag= a  
FT /product= "TbrA12"  
FT /partial  
FT /note= "No start codon"  
XX  
XX WO200124820-A1.  
XX  
XX 12-APR-2001.  
XX  
XX 10-OCT-2000; 2000WO-US28095.  
XX  
XX 07-OCT-1999; 99US-0158338.  
XX 07-OCT-1999; 99US-0158425.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;  
XX WPI; 2001-250576/30.  
DR P-PSDB; AAU01889.

XX Vaccinating against Mycobacterium tuberculosis in mammals using fusion  
 PT proteins comprising combinations of heterologous antigens -  
 PS Example 2; Page 162; 168pp; English.

XX The sequence encodes Mycobacterium tuberculosis TbrA12,  
 CC an M. tuberculosis antigen. Compositions comprising at least  
 CC 2 heterologous antigens, as a fusion protein, and vectors expressing the  
 CC fusion proteins are used as vaccines to prophylactically immunise  
 CC mammals (especially humans) against infection by Mycobacterium. The  
 CC compositions contain at least 2 heterologous antigens that increase the  
 CC serological sensitivity of individuals infected with tuberculosis, a  
 CC disease frequently affecting patients with acquired immunodeficiency  
 CC disease, AIDS.

XX Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:  
 Pred. No.: 8.81e-15 Length: 447  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-684-215A-17 (1-30) x AAS03780 (1-447)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 11 ACGGCCGGCTCGGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCCATTCGGATC 70  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 DB 71 GGGCAGGCGATGCGATCGCGGCCAGATC 100

RESULT 13

AAD47080  
 ID AAD47080 standard; DNA; 447 BP.

XX AAD47080;

XX 27-JAN-2003 (first entry)

XX Mycobacterium tuberculosis Ra12 antigen encoding DNA.

XX Vaccine; Immunity; diagnostic agent; gene therapy; Ra12 antigen;  
 KW gene; ds.

XX Mycobacterium tuberculosis.

XX Key Location/Qualifiers  
 FH CDS 11..409  
 FT /\*tag= a  
 FT /product= "Ra12 antigenic protein"  
 FT /note= "No start codon"  
 FT /partial

XX WO200272792-A2.

XX 19-SEP-2002.

XX 13-MAR-2002; 2002WO-US08223.

XX 13-MAR-2001; 2001US-275837P.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Brannon M, Guderian J;

XX WPI; 2002-759844/82.

XX P-PSDB; AAE29705.

XX

PT New recombinant nucleic acid molecule comprising a Leishmania TSA,  
 PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective  
 PT immunity against pathogenic microorganisms e.g. Leishmania and  
 PT Mycobacterium tuberculosis

XX Disclosure: Page 83; 155pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
 CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
 CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
 CC are used in methods for eliciting immune response in mammals. They are  
 CC useful as vaccines to elicit protective immunity against pathogenic  
 CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
 CC polypeptides are used for enhancing the expression of polynucleotides,  
 CC as in vivo diagnostic agents and for raising antibodies in a non-human  
 CC animal. The invention is used in gene therapy. The present sequence is  
 CC M. tuberculosis Ra12 (C-terminus of MTB32A; RA35FL) antigen encoding DNA.

XX Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:  
 Pred. No.: 8.81e-15 Length: 447  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-17 (1-30) x AAD47080 (1-447)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 11 ACGGCCGGCTCGGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCCATTCGGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

DB 71 GGGCAGGCGATGCGATCGCGGCCAGATC 100

RESULT 14

AAD28339  
 ID AAD28339 standard; cDNA; 447 BP.

XX AAD28339;

XX 22-APR-2002 (first entry)

XX Mycobacterium species MTB32A C-terminal peptide encoding cDNA, MTB32A.

XX Fusion protein; antigen; serological sensitivity; immune response;  
 KW tuberculosis; infection; vaccine; MTB32A; RA32FL; MTB32A; ss.

XX Mycobacterium sp.

XX Key Location/Qualifiers  
 FH CDS 11..409  
 FT /\*tag= a  
 FT /product= "Ra35 protein fragment"  
 FT /note= "CDS does not include start codon"  
 FT /partial

XX WO200198460-A2.

XX 27-DEC-2001.

XX 20-JUN-2001; 2001WO-US19959.

XX 20-JUN-2000; 2000US-0597796.

XX 01-FEB-2001; 2001US-265737P.

XX (CORI-) CORIXA CORP.



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:54:40 : Search time 191.628 Seconds  
(without alignments)  
1584.308 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TRASNDFQLSQGQGFAPFI.....QTKSGGTRGNTVLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool.p/US09684215/runat\_05092003\_072203\_619/app\_query.fasta\_1.853  
-DB=PublishedApplications\_NA\_QFMT-fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BL0SUM62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -CGEN=1.130 @runat\_05092003\_072203\_619  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LOGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	675	100.0	702	9	US-09-287-849-27 Sequence 27, Appl

2	675	100.0	702	12	US-10-359-460-27	Sequence 27, Appl
3	675	100.0	1068	10	US-09-712-363-15	Sequence 15, Appl
4	675	100.0	2287	9	US-09-287-849-1	Sequence 1, Appli
5	675	100.0	2287	12	US-10-359-460-1	Sequence 1, Appli
6	670	99.3	447	12	US-10-084-843-4	Sequence 4, Appli
7	670	99.3	447	12	US-10-193-002-4	Sequence 4, Appli
8	670	99.3	1872	12	US-10-084-843-17	Sequence 17, Appl
9	670	99.3	1872	12	US-10-193-002-17	Sequence 17, Appl
10	632	93.6	675	9	US-09-759-143-822	Sequence 822, App
11	632	93.6	675	9	US-09-780-669-822	Sequence 822, App
12	632	93.6	675	9	US-09-822-827-822	Sequence 822, App
13	632	93.6	675	10	US-09-895-793-822	Sequence 822, App
14	632	93.6	675	10	US-09-895-814-822	Sequence 822, App
15	632	93.6	675	12	US-10-144-678A-822	Sequence 822, App
16	632	93.6	675	13	US-10-012-896-822	Sequence 1862, Ap
17	632	93.6	822	10	US-09-736-457-1862	Sequence 1862, Ap
18	632	93.6	822	10	US-09-902-941-1862	Sequence 1862, Ap
19	632	93.6	822	10	US-09-849-626-1862	Sequence 1862, Ap
20	632	93.6	822	14	US-10-017-754-1862	Sequence 333, App
21	632	93.6	900	9	US-09-735-705-353	Sequence 353, App
22	632	93.6	900	10	US-09-850-716A-353	Sequence 353, App
23	632	93.6	900	10	US-09-897-778-353	Sequence 353, App
24	632	93.6	900	12	US-10-117-982-353	Sequence 353, App
25	632	93.6	915	9	US-09-759-143-834	Sequence 834, App
26	632	93.6	915	9	US-09-780-669-834	Sequence 834, App
27	632	93.6	915	9	US-09-822-827-834	Sequence 834, App
28	632	93.6	915	10	US-09-895-793-834	Sequence 834, App
29	632	93.6	915	10	US-09-895-814-834	Sequence 834, App
30	632	93.6	915	12	US-10-144-678A-834	Sequence 834, App
31	632	93.6	915	13	US-10-012-896-834	Sequence 1861, Ap
32	632	93.6	945	10	US-09-736-457-1861	Sequence 1861, Ap
33	632	93.6	945	10	US-09-902-941-1861	Sequence 1861, Ap
34	632	93.6	945	10	US-09-849-626-1861	Sequence 1861, Ap
35	632	93.6	945	14	US-10-017-754-1861	Sequence 351, App
36	632	93.6	1012	9	US-09-735-705-351	Sequence 351, App
37	632	93.6	1012	10	US-09-850-716A-351	Sequence 351, App
38	632	93.6	1012	10	US-09-897-778-351	Sequence 351, App
39	632	93.6	1012	12	US-10-117-982-351	Sequence 1084, Ap
40	632	93.6	1035	9	US-09-922-217-1084	Sequence 1084, Ap
41	632	93.6	1035	10	US-09-833-263-1084	Sequence 388, App
42	632	93.6	1035	11	US-09-938-864-388	Sequence 1084, Ap
43	632	93.6	1035	13	US-10-025-380-1084	Sequence 388, App
44	632	93.6	1035	14	US-10-125-635A-388	Sequence 388, App
45	632	93.6	1035	14	US-10-002-603-388	Sequence 388, App

#### ALIGNMENTS

#### RESULT 1

US-09-287-849-27

: Sequence 27, Application US/09287849

: Patent No. US2002009459A1

: GENERAL INFORMATION:

: APPLICANT: Reed, Steven G.

: APPLICANT: Skeiky, Yasir A.W.

: APPLICANT: Dillon, Davin C.

: APPLICANT: Alderson, Mark

: APPLICANT: Campos-Neto, Antonio

: APPLICANT: Corixa Corporation

: TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

: TITLE OF INVENTION: and Their Uses

: FILE REFERENCE: 014058-0090200S

: CURRENT APPLICATION NUMBER: US/09/287,849

: CURRENT FILING DATE: 1999-04-07

: PRIOR APPLICATION NUMBER: US 08/818,112

: PRIOR FILING DATE: 1997-03-13

: PRIOR APPLICATION NUMBER: US 08/942,578

: PRIOR FILING DATE: 1997-10-01

: PRIOR APPLICATION NUMBER: US 09/025,197

: PRIOR FILING DATE: 1998-02-18

: PRIOR APPLICATION NUMBER: US 09/056,556

: PRIOR FILING DATE: 1998-04-07

: PRIOR APPLICATION NUMBER: US 09/223,040

```

; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.: 2,266-69 Length: 702
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-287-849-27 (1-702)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 25 ACGGCGCGTCCGATACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGCGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
DB 85 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 144
QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60
DB 145 ATCGGGCCCTACCGCTTCCGCTTGGGTGTTCGACAAACACGCGGACGCGACGA 204
QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 265 ATCAACGCGGTGGTGGGAGCGCTCCGATCACTCGGCGGAGGATGGCGGACGCGTTAAC 324
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 325 GGGCATATATCCGGTGAGTGTATCTCGGTGACCTCGGCAACCAAGTCGGGGCGGACGCGT 384
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 385 ACAGGGAACGTGCATTTGGCGGAGGACCCCGGCC 420

RESULT 2
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

```

```

; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-0090200S
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-10-359-460-27

Alignment Scores:
Pred. No.: 2,266-69 Length: 702
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-4 (1-132) x US-10-359-460-27 (1-702)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 25 ACGGCGCGTCCGATACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGCGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
DB 85 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 144
QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60
DB 145 ATCGGGCCCTACCGCTTCCGCTTGGGTGTTCGACAAACACGCGGACGCGACGA 204
QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 205 GTCCAAACGCGTGGTGGGAGCGCTCCGCGCGCAAGTCTCGCATCTCCACCGCGACGCGT 264
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 265 ATCAACGCGGTGGTGGGAGCGCTCCGATCACTCGGCGGAGGATGGCGGACGCGTTAAC 324
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

```

Db 910 ATCACCGCGGTGCAGCGCGGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTTAAC 969

Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 |||||  
 Db 970 GGCATCATCCCGGTGACGTCTCTCGGTACCTGGCAACCAAGTCTGGGGGCGACCGGT 1029

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 |||||  
 Db 1030 ACAGGAAACGTGACATTGGCCGAGGACCCCGGCC 1065

RESULT 4

US-09-287-849-1  
 : Sequence 1, Application US/09287849  
 : Patent No. US20020009459A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Reed, Steven G.  
 : APPLICANT: Skeiky, Yasir A.W.  
 : APPLICANT: Dillon, Davin C.  
 : APPLICANT: Alderson, Mark  
 : APPLICANT: Campos-Neto, Antonio  
 : APPLICANT: Corixa Corporation  
 : TITLE OF INVENTION: Fusion proteins of Mycobacterium tuberculosis Antigens  
 : TITLE OF INVENTION: and Their Uses  
 : FILE REFERENCE: 014058-0090200S  
 : CURRENT APPLICATION NUMBER: US/09/287,849  
 : CURRENT FILING DATE: 1999-04-07  
 : PRIOR APPLICATION NUMBER: US 08/818,112  
 : PRIOR FILING DATE: 1997-03-13  
 : PRIOR APPLICATION NUMBER: US 08/942,578  
 : PRIOR FILING DATE: 1997-10-01  
 : PRIOR APPLICATION NUMBER: US 09/025,197  
 : PRIOR FILING DATE: 1998-02-18  
 : PRIOR APPLICATION NUMBER: US 09/056,556  
 : PRIOR FILING DATE: 1998-04-07  
 : PRIOR APPLICATION NUMBER: US 09/223,040  
 : PRIOR FILING DATE: 1998-12-30  
 : NUMBER OF SEQ ID NOS: 46  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 1  
 : LENGTH: 2287  
 : TYPE: DNA  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
 : OTHER INFORMATION: protein Ral2-TbH9-Ra35 (designated Mtb32-Mtb39  
 : OTHER INFORMATION: fusion)  
 : NAME/KEY: modified\_base  
 : LOCATION: (30)  
 : OTHER INFORMATION: n = g, a, c or t  
 : NAME/KEY: modified\_base  
 : LOCATION: (33)  
 : OTHER INFORMATION: n = g, a, c or t  
 : NAME/KEY: CDS  
 : LOCATION: (42)..  
 : NAME/KEY: modified\_base  
 : LOCATION: (2270)  
 : OTHER INFORMATION: n = g, a, c or t  
 : US-09-287-849-1

Alignment Scores:  
 Pred. No.: 9,31e-69 Length: 2287  
 Score: -675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-287-849-1 (1-2287)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 |||||  
 Db 63 ACGGCCGCGTCCGATAACTTCCAGCTCTCCAGGGTGGCAGGATTCGGCATTCGCATC 122

```
OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 123 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCAACCCACCGTTCAT 182
OY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 183 ATCGGGCCCTACCGCCCTCTCGGCTGGGTGGTGTGTGTCGACAAACGCAACGGCGCACGA 242
OY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCACACGGTGGTGGGAGCGCTCGGGCGGCAAGTCTCGGATCTCCACCGCGCACGTG 302
OY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCGCGTGGCGCGCTCGGATCAACTCGGCGCACCGATGCGGCGCACGCGTTCAT 362
OY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 363 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGCACGCGT 422
OY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 423 ACAGGAACGTGACATTGGCCGAGGACCCCGGCC 458
```

## RESULT 5

```
US-10-359-460-1
; Sequence 1, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR FILING DATE: 1997-10-01
; PRIOR FILING DATE: 1999-04-07
; PRIOR FILING DATE: 1997-03-13
; PRIOR FILING DATE: 1997-03-13
; PRIOR FILING DATE: 1997-10-01
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-04-07
; PRIOR FILING DATE: 1998-04-07
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mt32-Mtb39
; OTHER INFORMATION: fusion)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
```

```
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-10-359-460-1
Alignment Scores:
Pred. No.: 9,31e-69 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-09-684-215A-4 (1-132) x US-10-359-460-1 (1-2287)
OY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 63 ACAGCGCGGTTCGATAAATCCAGCTGTCCAGGGTGGCGAGGGATTCGCCATTCGGATC 122
OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 123 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCAACCCACCGTTCAT 182
OY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 183 ATCGGGCCCTACCGCCCTCTCGGCTGGGTGGTGTGTGTCGACAAACGCAACGGCGCACGA 242
OY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCACACGGTGGTGGGAGCGCTCGGGCGGCAAGTCTCGGATCTCCACCGCGCACGTG 302
OY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCGCGTGGCGCGCTCGGATCAACTCGGCGCACCGATGCGGCGCACGCGTTCAT 362
OY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 363 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGCACGCGT 422
OY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 423 ACAGGAACGTGACATTGGCCGAGGACCCCGGCC 458
```

## RESULT 6

```
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```



```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
;
; US-10-084-843-4
;
; Alignment Scores:
; Pred. No.: 5,01e-69 Length: 447
; Score: 670.00 Matches: 131
; Percent Similarity: 99.24% Conservative: 0
; Best Local Similarity: 99.24% Mismatches: 1
; Query Match: 99.26% Indels: 0
; DB: 12 Gaps: 0
;
; US-09-684-215A-4 (1-132) x US-10-084-843-4 (1-447)
;
; QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
; Db 11 ACGGCGCGCTCCGATAAATCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGCATC 70
;
; QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
; Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTCAT 130
;
; QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60
; Db 131 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCACAAACGCGACGCGCTAAC 190
;
; QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
; Db 191 GTCCAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
;
; QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
; Db 251 ATCACGCGGTGACGCGCTCCGATCAACTCGGCCACCGCATCCACCGCGCTAAC 310
;
; QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
; Db 311 GGGCATCATCCCGTACGCTCATCTCGGTGAACCTGGCAACCAAGTCGGGCGGACGCGT 370
;
; RESULT 7
; US-10-193-002-4
; Sequence 4, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skelky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonia
; Houghton, Raymond
; Vedrick, Thomas S.
; Twardzik, Daniel R.
;
; US-09-684-215A-4 (1-132) x US-10-193-002-4 (1-447)
;
; QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
; Db 11 ACGGCGCGCTCCGATAAATCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGCATC 70
;
; QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
; Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTCAT 130
;
; QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60
; Db 131 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCACAAACGCGACGCGCTAAC 190
;
; QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
; Db 191 GTCCAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
;
; QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
; Db 251 ATCACGCGGTGACGCGCTCCGATCAACTCGGCCACCGCATCCACCGCGCTAAC 310
;
; QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
; Db 311 GGGCATCATCCCGTACGCTCATCTCGGTGAACCTGGCAACCAAGTCGGGCGGACGCGT 370
;
; Alignment Scores:
; Pred. No.: 5,01e-69 Length: 447
; Score: 670.00 Matches: 131
; Percent Similarity: 99.24% Conservative: 0
; Best Local Similarity: 99.24% Mismatches: 1
; Query Match: 99.26% Indels: 0
; DB: 12 Gaps: 0
;
; US-09-684-215A-4 (1-132) x US-10-193-002-4 (1-447)
;
; QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
; Db 11 ACGGCGCGCTCCGATAAATCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGCATC 70
;
; QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
; Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTCAT 130
;
; QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60
; Db 131 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCACAAACGCGACGCGCTAAC 190
;
; QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
; Db 191 GTCCAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
;
; QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
; Db 251 ATCACGCGGTGACGCGCTCCGATCAACTCGGCCACCGCATCCACCGCGCTAAC 310
;
; QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
; Db 311 GGGCATCATCCCGTACGCTCATCTCGGTGAACCTGGCAACCAAGTCGGGCGGACGCGT 370
;
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/193,002
; FILING DATE: 10-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
;
; US-10-193-002-4
;
; Alignment Scores:
; Pred. No.: 5,01e-69 Length: 447
; Score: 670.00 Matches: 131
; Percent Similarity: 99.24% Conservative: 0
; Best Local Similarity: 99.24% Mismatches: 1
; Query Match: 99.26% Indels: 0
; DB: 12 Gaps: 0
;
; US-09-684-215A-4 (1-132) x US-10-193-002-4 (1-447)
;
; QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
; Db 11 ACGGCGCGCTCCGATAAATCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGCATC 70
;
; QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
; Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTCAT 130
;
; QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60
; Db 131 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCACAAACGCGACGCGCTAAC 190
;
; QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
; Db 191 GTCCAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
;
; QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
; Db 251 ATCACGCGGTGACGCGCTCCGATCAACTCGGCCACCGCATCCACCGCGCTAAC 310
;
; QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
; Db 311 GGGCATCATCCCGTACGCTCATCTCGGTGAACCTGGCAACCAAGTCGGGCGGACGCGT 370
```



Alignment Scores:  
 Pred. No.: 2,79e-68 Length: 1872  
 Score: 670.00 Matches: 131  
 Percent Similarity: 99.24% Conservative: 0  
 Best Local Similarity: 99.24% Mismatches: 1  
 Query Match: 99.26% Indels: 0  
 DB: 12 Gaps: 0

US-09-684-215A-4 (1-132) x US-10-193-002-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 758 ACGCCGCGTCCGATAACTTCAGCTGTCCAGGGTGGCAGGATTGCCATTCCGATC 817  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 818 GGGCAGCGATGGCGATCGCGGCAATCCGATCGGGTGGGGTCAACCCCGTTCAT 877  
 Qy 41 IleGlyProThrAlaPheLeuGlyLeuValValAlaAspAsnAsnGlyAlaAlaArg 60  
 Db 878 ATCGGGCTACCGCTTCCTCGCTTGGTGTTCGACAAACAACGGCAACGGCGCAGCA 937  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 938 GTCCAACGCGTGGTGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCAGTG 997  
 Qy 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsn 100  
 Db 998 ATCAGCGGGTGCAGGGCTCCGATCAATCGGCCACCGCATGGCGGAGCGCGTTAAC 1057  
 Qy 101 GlyHisHsProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 1058 GGGCATCATCCGCTGAGCTCATCTCGGTGAATCGGCAACCAAGTCGGGGCGCAGCGGT 1117  
 Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 1153

RESULT 10

US-09-759-143-822  
 ; Sequence 822, Application US/09759143  
 ; Patent No. US2002002248A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, David C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yuqi  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Li, Samuel  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Hepler, William  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.427C23  
 ; CURRENT APPLICATION NUMBER: US/09/759,143  
 ; CURRENT FILING DATE: 2001-01-12  
 ; NUMBER OF SEQ ID NOS: 934  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 822  
 ; LENGTH: 675  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-759-143-822

Alignment Scores:  
 Pred. No.: 2.1e-64 Length: 675  
 Score: 632.00 Matches: 126  
 Percent Similarity: 96.21% Conservative: 1  
 Best Local Similarity: 95.45% Mismatches: 1  
 Query Match: 93.63% Indels: 4  
 DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-759-143-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 22 ACGCCGCGTCCGATAACTTCAGCTGTCCAGGGTGGCAGGATTGCCATTCCGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 82 GGGCAGCGATGGCGATCGCGGCGCAGATCAATC-----CTTCCCGCGTTCAT 129  
 Qy 41 IleGlyProThrAlaPheLeuGlyLeuValValAlaAspAsnAsnGlyAlaAlaArg 60  
 Db 130 ATCGGGCTACCGCTTCCTCGCTTGGTGTTCGACAAACAACGGCAACGGCGCAGCA 189  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 190 GTCCAACGCGTGGTGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCAGTG 249  
 Qy 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsn 100  
 Db\* 250 ATCAGCGGGTGCAGGGCTCCGATCAATCGGCCACCGCATGGCGGAGCGCGTTAAC 309  
 Qy 101 GlyHisHsProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 310 GGGCATCATCCGCTGAGCTCATCTCGGTGAATCGGCAACCAAGTCGGGGCGCAGCGGT 369  
 Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 405

RESULT 11

US-09-780-669-822  
 ; Sequence 822, Application US/09780669  
 ; Patent No. US2002005197A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, David C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yuqi  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalcs, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Li, Samuel  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Hepler, William  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Houghton, Raymond L.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.427C24  
 ; CURRENT APPLICATION NUMBER: US/09/780,669  
 ; CURRENT FILING DATE: 2001-02-09  
 ; NUMBER OF SEQ ID NOS: 943  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 822  
 ; LENGTH: 675

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-780-669-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 40
Db 82 GGGCAGGCGATCGCGATCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 129
QY 41 IleGlyProThrAlaPheLeuGlyGlnIleArgSerGlyGlySerProThrValHls 60
Db 130 ATCGGGCCCTACCGCTTCTCGCTTGGGTGTGTCGACAAACGCGGACGCGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATCACCAGCGGTGGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAC 309
QY 101 GlyHlsHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCCGGTGAGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACG 369

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 370 ACAGGGAACGTGACATTGGCCGAGGAGGCCCGCGGCC 405

RESULT 13
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 10 Gaps: 1

US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-822-827-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 40
Db 82 GGGCAGGCGATCGCGATCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 129
QY 41 IleGlyProThrAlaPheLeuGlyGlnIleArgSerGlyGlySerProThrValHls 60
Db 130 ATCGGGCCCTACCGCTTCTCGCTTGGGTGTGTCGACAAACGCGGACGCGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATCACCAGCGGTGGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAC 309
QY 101 GlyHlsHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCCGGTGAGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACG 369

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 370 ACAGGGAACGTGACATTGGCCGAGGAGGCCCGCGGCC 405

RESULT 12
US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-822-827-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 40
Db 82 GGGCAGGCGATCGCGATCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 129
QY 41 IleGlyProThrAlaPheLeuGlyGlnIleArgSerGlyGlySerProThrValHls 60
Db 130 ATCGGGCCCTACCGCTTCTCGCTTGGGTGTGTCGACAAACGCGGACGCGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATCACCAGCGGTGGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAC 309
QY 101 GlyHlsHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCCGGTGAGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACG 369

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 370 ACAGGGAACGTGACATTGGCCGAGGAGGCCCGCGGCC 405
```

US-09-684-215A-4 (1-132) x US-09-895-793-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACCGCCGCTCCAGTTCACAGCTTCAGGTTGGGAGGATTCGCATTCGCATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 Db 82 GGGCAGCGATGCGATCGGGGCCAGATCAAG-----CTTCCACCGTTTCAT 129  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
 Db 130 ATCCAGCGCTACCCCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGACGCGACGA 189  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 190 GTCCACGCGTGTGCGGAGCGTCCGCGGCAAGTCTCGCATCTCCACCGCGAGGTG 249  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 250 ATCCAGCGGTCGACGCGCTCGGATCAACTCGGCCACCGCATGGCGGACGCGCTAAC 309  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArg 120  
 Db 310 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 369  
 QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
 Db 370 ACAGGACGTGACATGGCGGAGGACCCCGGCC 405

RESULT 14

US-09-895-814-822  
 ; Sequence 822, Application US/09895814  
 ; Publication No. US20020193296A1  
 ; GENERAL INFORMATION:

; APPLICANT: Xu, JIANGCHUN  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yuqiu  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Retter, Gary W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aljun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Hural, John  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Houghton, Raymond L.  
 ; APPLICANT: Vinals de Bassols, Carlota  
 ; APPLICANT: Foy, Teresa  
 ; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.427C26  
 ; CURRENT APPLICATION NUMBER: US/09/895,814  
 ; CURRENT FILING DATE: 2001-06-29  
 ; NUMBER OF SEQ ID NOS: 990  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 822

; LENGTH: 675  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-895-814-822

Alignment Scores: 2.1e-64 Length: 675  
 Pred. No.:

Score: 632.00 Matches: 126  
 Percent Similarity: 96.21% Conservative: 1  
 Best Local Similarity: 95.45% Mismatches: 1  
 Query Match: 93.63% Indels: 4  
 DB: 10 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-895-814-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACCGCCGCTCCAGTTCACAGCTTCAGGTTGGGAGGATTCGCATTCGCATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 Db 82 GGGCAGCGATGCGATCGGGGCCAGATCAAG-----CTTCCACCGTTTCAT 129  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
 Db 130 ATCCAGCGCTACCCCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGACGCGACGA 189  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 190 GTCCACGCGTGTGCGGAGCGTCCGCGGCAAGTCTCGCATCTCCACCGCGAGGTG 249  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 250 ATCCAGCGGTCGACGCGCTCGGATCAACTCGGCCACCGCATGGCGGACGCGCTAAC 309  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArg 120  
 Db 310 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 369  
 QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
 Db 370 ACAGGACGTGACATGGCGGAGGACCCCGGCC 405

RESULT 15

US-10-144-678A-822  
 ; Sequence 822, Application US/10144678A  
 ; Publication No. US20030157089A1  
 ; GENERAL INFORMATION:

; APPLICANT: Xu, JIANGCHUN  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yuqiu  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aljun  
 ; APPLICANT: Skeiky, Yasir A. W.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Hural, John  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Houghton, Raymond L.  
 ; APPLICANT: Vinals y de Bassols, Carlota  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: Watanabe, Yoshihiro  
 ; APPLICANT: Deng, Ta

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.427C28  
 ; CURRENT APPLICATION NUMBER: US/10/144,678A  
 ; CURRENT FILING DATE: 2002-08-12  
 ; NUMBER OF SEQ ID NOS: 1033  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 822

; LENGTH: 675  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-144-678A-822

Alignment Scores:  
 Pred. No.: 2.le-64 Length: 675  
 Score: 632.00 Matches: 126  
 Percent Similarity: 96.21% Conservative: 1  
 Best Local Similarity: 95.45% Mismatches: 1  
 Query Match: 93.63% Indels: 4  
 DB: 12 Gaps: 1

US-09-684-215A-4 (1-132) x US-10-144-678A-822 (1-675)

QY	1	ThrAlaAlaSerAspAsn	PheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle	20
DB	22	ACGGCGCGTCCGATCACT	TCCAGCTGCCCCAGGTGGCAGGGATTCCCATTCGGATC	81
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40	
DB	82	GGGCAGGCGATGGCGATCGCGGCCAGATCAAG	-----CTTCCCACCGTTTCAT	129
QY	41	IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg	60	
DB	130	ATCGGGCCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCGCACGA	189	
QY	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80	
DB	190	GTCCAAACGGTGGTCCGGAGCGCTCCGGGGGCAAGTCTCGGCATCTCCACCGGCACGTG	249	
QY	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100	
DB	250	ATCACCGCGTCCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC	309	
QY	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg	120	
DB	310	GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGCACGCT	369	
QY	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132	
DB	370	ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC	405	

Search completed: September 5, 2003, 12:04:43  
 Job time : 196.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:24:34 ; Search time 248.497 Seconds  
(without alignments)  
1390.474 Million cell updates/sec

Title: US-09-684-215A-18

Perfect score: 653

Sequence: 1 TRAASNFLSQSGGQFAIP1.....OTKSGGTRGNTVLRGPPA 128

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cqn2\_1/USPTO.spool\_p/US09684215/runat\_05092003\_072201\_548/app\_query.fasta\_1.853  
-DB=N\_Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09684215 -CGN\_1.1\_526 @runat\_05092003\_072201\_548 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	653	100.0	675	22	AAS64132	Human /M. tubercul
2	653	100.0	675	22	AH93896	Ral2-P510S-C const
3	653	100.0	675	24	ABL95503	Ral2- P510S-C cons
4	653	100.0	675	25	ACA59940	Prostate cancer th
5	653	100.0	822	24	ABK39769	DNA encoding lung
6	653	100.0	822	25	ACA12098	Human lung cancer
7	653	100.0	822	25	ACA03284	Lung cancer therap
8	653	100.0	894	24	ABK27798	Human colon cancer
9	653	100.0	900	24	ABO92443	Human lung cancer
10	653	100.0	900	24	ABL49257	Ral2/C-terminal po
11	653	100.0	915	22	AAS64141	Human /M. tubercul
12	653	100.0	915	22	AH93905	Ral2-P775P-ORF3 c
13	653	100.0	915	24	ABL95512	Ral2- P775P-ORF3 c
14	653	100.0	915	25	ACA59949	Prostate cancer th
15	653	100.0	945	24	ABK39768	DNA encoding lung
16	653	100.0	945	25	ACA12097	Human lung cancer
17	653	100.0	945	25	ACA03283	Lung cancer therap
18	653	100.0	1012	24	ABO92442	Human lung cancer
19	653	100.0	1012	24	ABL49256	Ral2/N-terminal po
20	653	100.0	1035	24	ABK69714	DNA encoding human
21	653	100.0	1035	25	ABK33699	Human colon tumour
22	653	100.0	1155	24	ABK39775	DNA encoding Ral2-
23	653	100.0	1155	25	ACA12104	Human Ral2-l985P f
24	653	100.0	1155	25	ACA03290	Lung cancer therap
25	653	100.0	1203	22	AAS64153	Human /M. tubercul
26	653	100.0	1203	22	AH93917	Ral2-P501S-E2 cons
27	653	100.0	1203	24	ABL95524	Ral2-P501S-E2 cons
28	653	100.0	1203	25	ACA59961	Prostate cancer th
29	653	100.0	1263	24	ABK69715	DNA encoding human
30	653	100.0	1464	22	AH56353	DNA encoding Chlam
31	653	100.0	1464	24	ABL92582	Chlamydia trachoma
32	653	100.0	1557	22	AH56341	DNA encoding Chlam
33	653	100.0	1557	24	ABL92570	Chlamydia trachoma
34	653	100.0	1578	21	AA64764	C. pneumoniae sero
35	653	100.0	1578	22	AH56267	Chlamydia trachoma
36	653	100.0	1578	24	ABU92496	Chlamydia DNA sequ
37	653	100.0	1590	24	ABU99869	Breast tumour-spec
38	653	100.0	1707	24	ABK69716	DNA encoding human
39	653	100.0	1752	22	AH56356	DNA encoding Chlam
40	653	100.0	1752	24	ABL92585	Chlamydia trachoma
41	653	100.0	1758	22	AH56344	DNA encoding Chlam
42	653	100.0	1758	24	ABU92573	Chlamydia trachoma
43	653	100.0	1758	25	ACA59997	Prostate cancer th
44	653	100.0	1860	22	AH56323	DNA encoding Chlam
45	653	100.0	1860	24	ABL92552	Chlamydia trachoma

# ALIGNMENTS

RESULT 1

AAS64132

ID AAS64132 standard; cDNA; 675 BP.

XX AAS64132;

XX 29-JAN-2002 (first entry)

DE Human /M. tuberculosis Ral2 fusion protein Ral2-P510S-C cDNA.

XX Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.

XX Chimeric - Homo sapiens.

OS Chimeric - Microbacterium tuberculosis.

OS Synthetic.

PN WO200173032-A2.

```
XX PD 04-OCT-2001.
XX PF 27-MAR-2001; 2001WO-US09919.
XX PR 27-MAR-2000; 2000US-0536857.
XX PR 09-MAY-2000; 2000US-0568100.
XX PR 12-MAY-2000; 2000US-0570737.
XX PR 13-JUN-2000; 2000US-0593793.
XX PR 27-JUN-2000; 2000US-0605783.
XX PR 10-AUG-2000; 2000US-0636215.
XX PR 29-AUG-2000; 2000US-0651236.
XX PR 06-SEP-2000; 2000US-0657279.
XX PR 02-OCT-2000; 2000US-0679426.
XX PR 10-OCT-2000; 2000US-0685166.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stoik JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX WPI: 2001-639232/73.
XX DR P-PSDB: AAU69899.
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
XX the diagnosis and treatment of cancer, especially prostate cancer -
XX Example 17; Page 532; 579pp; English.
XX
XX The invention relates to isolated prostate-specific
XX polynucleotides, polypeptides, fusion proteins of the polypeptides,
XX antibodies raised against the polypeptides (or antigenic epitopes
XX derived from them) and antigen-presenting cells expressing the
XX polypeptides. The antibodies are useful for detecting the presence of
XX cancer, especially prostate cancer. The polypeptides, polynucleotides and
XX the antigen-presenting cells are useful for stimulating and/or expanding
XX T cells specific for a tumour protein, and for inhibiting the development
XX of cancer especially prostate cancer. Compositions comprising the
XX polynucleotide and/or polypeptide are useful for stimulating an immune
XX response, and for treating cancer. The oligonucleotide is useful for
XX detecting cancer. The present sequence is a prostate specific
XX polynucleotide of the invention.
XX
XX Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 5, 71e-58 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-684-215a-18 (1-128) x AAS64132 (1-675)
QY 1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleIleLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGGCGATGGCGATCGGGGCGAGATCAAGCTCCACCGTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGCTTGGGTGTGTCACACACACGACGCGGCGGACGATCCAGCGGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGGACGTATCACCGCGTTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100

Db 262 GACGGCGCTCGCATCACTCGCCACCGCATGGCGGACGGCTTAACGGCGCATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACCGTACAGGGANCGTG 381
QY 121 ThrLeuAlaGlnGlyProProAla 128
DB 382 ACATTGGCGGAGGGACCCCGGCC 405

RESULT 2
AAH93896
ID AAH93896 standard; cDNA; 675 BP.
XX AC AAH93896;
XX DT 04-OCT-2001 (first entry)
XX DE Ra12-PS10S-C construct cDNA sequence.
XX KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
XX KW cytostatic; gene therapy; metastasis; ss.
XX OS Homo sapiens.
XX PN WO200151633-A2.
XX PD 19-JUL-2001.
XX PF 16-JAN-2001; 2001WO-US01574.
XX PR 14-JAN-2000; 2000US-0483672.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stoik JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX DR WPI: 2001-425873/45;
XX PT New polynucleotide encoding a prostate-specific protein, for
XX PT diagnosing, monitoring and treating prostate cancer in a patient and
XX PT for use in vaccines -
XX PS Claim 8; Page 492-493; 543pp; English.
XX CC The present invention describes polynucleotide sequences (I) which encode
XX CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
XX CC and can be used in vaccine production and gene therapy. (I), (II),
XX CC antibodies to (II), fusion proteins comprising (II), and isolated
XX CC T cells prepared using (I) or (II) are used treat cancer in a patient.
XX CC (I) and the antibodies are also used in the detection of cancer in a
XX CC patient. The cancer that is diagnosed or treated is particularly
XX CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
XX CC (I) can be used for monitoring the progression of cancer in a patient.
XX CC (I) and (II) can also be used to improve diagnostic and therapeutic
XX CC methods for prostate cancer. They can indicate the level of metastasis
XX CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
XX CC AAH01318 represent polynucleotide and amino acid sequences used in the
XX CC exemplification of the present invention.
XX SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 5, 71e-58 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
```



US-09-684-215A-18 (1-128) x AAH93896 (1-675)

```
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20
Db 22 ACGCGCGCTCCGATACTCCAGCTGTCAGAGGTGGCAGGATTCCGATTCGGATC 81
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGCATGGCGATCGCGGCCAGATCAAGCTCCACCGTTTCATATCGGCGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGCTGGGTGGTGTGTCGACAAACGCGACGCGGATCCACGCGGTC 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGATGATCACGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCTCCGATCAACTCGGCCACCGCATGCGGCGGCGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC. 405
RESULT 3
ABL95503
ID ABL95503 standard; cDNA; 675 BP.
AC ABL95503;
XX
DT 19-JUL-2002 (first entry)
XX
DE Ra12- P510S-C construct cDNA sequence SEQ ID NO 822.
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy; gene; ss.
OS Chimeric - Mycobacterium tuberculosis.
OS Chimeric - Homo sapiens.
XX
PN US2002022248-A1.
XX
PD 21-FEB-2002.
XX
PF 12-JAN-2001; 2001US-0759143.
XX
PR 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1998; 98US-0030607.
PR 14-JUL-1998; 98US-0115453.
PR 23-SEP-1998; 98US-0159812.
PR 15-JAN-1999; 99US-0232149.
PR 09-APR-1999; 99US-0288946.
PR 13-JUL-1999; 99US-0352616.
PR 18-NOV-1999; 99US-0439313.
PR 14-JAN-2000; 2000US-0483672.
PR 27-MAR-2000; 2000US-0536857.
PR 09-MAY-2000; 2000US-0568100.
PR 12-MAY-2000; 2000US-0570737.
PR 13-JUN-2000; 2000US-0593793.
PR 27-JUN-2000; 2000US-0605783.
PR 10-AUG-2000; 2000US-0636215.
PR 29-AUG-2000; 2000US-0651236.
PR 06-SEP-2000; 2000US-0657279.
PR 02-OCT-2000; 2000US-0679426.
```

10-OCT-2000; 2000US-0685166.

```
PR XX XUJ/J/ XU J.
PA PA DILL/ DILLON D C.
PA PA MITC/ MITCHAM J L.
PA PA HARL/ HARLOCKER S L.
PA PA JIAN/ JIANG Y.
PA PA KALO/ KALOS M D.
PA PA FANG/ FANGER G R.
PA PA RETT/ RETTER M W.
PA PA STOL/ STOLK J A.
PA PA DAYC/ DAY C H.
PA PA VEDV/ VEDVICK T S.
PA PA CART/ CARTER D.
PA PA LISX/ LI S X.
PA PA WANG/ WANG A.
PA PA SKEI/ SKEIKY Y A W.
PA PA HEPL/ HEPLER W T.
PA PA HEND/ HENDERSON R A.
XX XX
PI PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX XX
DR WPI; 2002-255649/30.
XX XX
PT New prostate-specific polynucleotides for diagnosing and treating
PT diseases, in particular prostate cancer, and as markers for the
PT progression of cancer -
XX XX
PS Example 17; SEQ ID NO 822; 87pp; English.
XX XX
CC The present invention provides prostate-specific coding sequences and
CC their encoded proteins. These can be used in the diagnosis and treatment
CC of cancers, particularly prostate cancer. The present sequence is a cDNA
CC described in the invention.
XX XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;
```

#### Alignment Scores:

Pred. No.:	5,71e-58	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-684-215A-18 (1-128) x ABL95503 (1-675)

```
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20
Db 22 ACGCGCGCTCCGATACTCCAGCTGTCAGAGGTGGCAGGATTCCGATTCGGATC 81
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGCATGGCGATCGCGGCCAGATCAAGCTCCACCGTTTCATATCGGCGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGCTGGGTGGTGTGTCGACAAACGCGACGCGGATCCACGCGGTC 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGGATGATCACGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCTCCGATCAACTCGGCCACCGCATGCGGCGGCGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGTACAGGGAACGTG 381
```

QY 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 4

ACA59940  
 ID ACA59940 standard; cDNA; 675 BP.  
 AC ACA59940;

XX 10-JUN-2003 (first entry)

XX Prostate cancer therapy associated cDNA #647.

XX Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA;  
 KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 KW PSMA; gene; ss.

OS Homo sapiens.

XX US2002192763-A1.

XX 19-DEC-2002.

XX 29-JUN-2001; 2001US-0895793.

XX 17-APR-2000; 2000US-157455P.

XX 04-OCT-2000; 2000US-0679272.

XX 28-MAR-2001; 2001US-0822827.

XX (XUJ/J) XU J.

PA (DILL/) DILLON D C.

PA (MITC/) MITCHAM J L.

PA (HARL/) HARLOCKER S L.

PA (JIAN/) JIANG Y.

PA (KALO/) KALOS M D.

PA (FANG/) FANGER G R.

PA (RETT/) RETTER M W.

PA (STOL/) STOLK J A.

PA (DAYC/) DAY C H.

PA (VEDV/) VEDVICK T S.

PA (CART/) CARTER D.

PA (LISX/) LI S X.

PA (WANG/) WANG A.

PA (SKEI/) SKEIKY Y A W.

PA (HEPL/) HEPLER W T.

PA (HEND/) HENDERSON R A.

PA (HURA/) HURAL J.

PA (MCNE/) MCNEILL P D.

PA (HOUG/) HOUGHTON R L.

PA (DBAS/) Y DE BASSOLS C V.

PA (FOVT/) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;

PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;

PI McNeill PD, Houghton RL, Y De Bassols CV, Foy TW;

DR WPI; 2003-352711/33.

XX New fusion protein comprising prostate-specific polypeptides, or its

PT immunogenic portions, useful for diagnosing, preventing and/or treating

PT cancer, particularly prostate cancer

XX Example 17; SEQ ID NO 822; 85pp; English.

PS The invention describes a fusion protein comprising at least one amino

XX acid sequence of immunogenic portions of any of the 3 sequences not

CC defined in the specification, or sequences having at least 70 or 90 %

CC sequence identity to any one of the 35 sequences defined in the USPTO

CC web site, which is encoded by any of the 4 nucleotide sequences not

CC defined in the specification. The fusion protein, composition and

CC methods are useful for diagnosing, preventing and/or treating cancer,  
 CC particularly prostate cancer. The proteins are useful as markers to  
 CC indicate the presence or absence of cancer. This sequence  
 CC represents a prostate cancer therapy associated cDNA.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=US20020192763.

XX Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;

Alignment Scores:

Pred. No.: 5,71e-58 Length: 675  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0

US-09-684-215A-18 (1-128) x ACA59940 (1-675)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACAGCGCGGTGCGATAACTTCCAGCTGTCACAGGTGGCAGGGATTCCCATTCGGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCCTCGGCTTGGGTGTTCTGACACAAACGACGCGGACGAGTCAACGCGGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGGTCCGGCGGCAAGTCTCGGCATCTCCACGGGCGAGTGTACCGGGTCC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGTCTCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAACGGGCATCATCC 321  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGTACAGGGAACGTC 381  
 QY 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 5

ABK39769  
 ID ABK39769 standard; cDNA; 822 BP.

XX AC ABK39769;

XX DT 21-MAY-2002 (first entry)

XX DNA encoding lung tumour protein P801P ORF5 and Ra12 fusion protein.

XX Lung tumour; cancer; T cell; Immune response stimulator;

XX cytosolic; gene; ss.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200204514-A2.

XX PD 17-JAN-2002.

XX PF 10-JUL-2001; 2001WO-US22058.

XX PR 11-JUL-2000; 2000US-0614124.

XX PR 29-AUG-2000; 2000US-0651563.

PR 08-SEP-2000; 2000US-0658824.  
PR 26-SEP-2000; 2000US-0671325.  
PR 06-OCT-2000; 2000US-0677419.  
PR 30-OCT-2000; 2000US-0702705.  
PR 13-DEC-2000; 2000US-0736457.  
PR 03-MAY-2001; 2001US-0849626.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;  
PI Marnerakis M, Carter D, Fanger GR, Vedwick TS, Bangur CS;  
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
XX  
DR WPI: 2002-164634/21.  
DR P-PSDB; AAU85587.  
XX  
PT Novel polynucleotide encoding a lung tumour polypeptide useful for  
PT stimulating and/or expanding T cells specific for a tumour protein  
XX  
PS Example 8; SEQ ID No 1862; 223pp; English.  
XX  
XX The invention describes an isolated polynucleotide and polypeptide  
CC useful for stimulating and/or expanding T cells specific for a tumour  
CC protein for determining the presence of a cancer in a patient. A  
CC composition containing the polynucleotide and/or polypeptide is useful  
CC for treating a lung cancer in a patient. The polypeptide is useful for  
CC removing tumour cells from a biological sample. The polynucleotide is  
CC also useful as probe or primer to detect the level of mRNA encoding a  
CC tumour protein. This sequence encodes a lung tumour associated protein  
CC or protein fragment, described in the method of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 other;

Alignment Scores:  
Pred. No.: 7,23e-58 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABK39769 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGCATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGCGATGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGCGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCCTCGGCTGGGTGTGTGTCGACAAACAGCGCAACGCGCAGTCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACGTGATCACCGGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGCTCCGATCAACTCGCGCCACCGCGATGGCGACGCGCTTAACGGGCATCCTCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 121 ThrLeuAlaGluGlyProProAla 128

Db 382 ACATTGCGCGAGGAGGACCCCGGCC 405  
RESULT 6  
ACAL2098  
XX ID ACAL2098 standard; cDNA; 822 BP.  
XX ACAL2098;  
XX  
XX 06-JUN-2003 (first entry)  
XX Human lung cancer protein L801P ORF5/Ra12 fusion protein cDNA.  
XX  
XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
KW T cell expansion; CD4; CD8; RA12; gene.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX US2002197669-A1.  
XX  
XX 26-DEC-2002.  
XX  
XX 03-MAY-2001; 2001US-0849626.  
XX  
XX 13-DEC-2000; 2000US-0736457.  
XX (BANG/) BANGUR C S.  
PA (FANG/) FANGER G R.  
PA (WANG/) WANG A.  
PA (WANG/) WANG T.  
PA (SWIT/) SWITZER A P.  
PA (MCNEI/) MCNEILL P D.  
PA (CLAP/) CLAPPER J D.  
XX  
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
PI Clapper JD;  
XX  
XX WPI: 2003-352750/33.  
DR P-PSDB; ABU69562.  
XX  
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful  
PT for detecting the presence of lung cancer in a patient, and in  
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer  
XX  
XX Example 8; Page -: 72pp; English.  
PS  
XX The invention relates to a polynucleotide encoding a lung tumour protein,  
CC comprising a sequence selected from any of the 14 sequences  
CC mentioned in the specification, or a sequence (S2) mentioned in  
CC specification, complement of S1, sequences consisting of at least 20  
CC contiguous residues of S1, sequences that hybridise to S1, sequences  
CC having 75%, preferably 90%, identity to S1, or degenerate variants of  
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)  
CC selected from any one of the 4 amino acid sequences mentioned in the  
CC specification, a sequence encoded by the polynucleotide, or sequences  
CC having at least 70%, preferably 90%, identity to a sequence encoded by  
CC the polynucleotide), an expression vector comprising the polynucleotide  
CC operably linked to an expression control sequence, a host cell  
CC transformed or transfected with the vector, an isolated antibody (or its  
CC antigen-binding fragment) that specifically binds to the polypeptide,  
CC detecting the presence of a cancer in a patient, a fusion protein  
CC comprising the polypeptide, an oligonucleotide that hybridises to  
CC S1 under moderately stringent conditions, stimulating and/or expanding T  
CC cells specific for a tumour protein (comprising contacting T cells with  
CC the polynucleotide, protein or antigen-presenting cells, under conditions  
CC and for a time sufficient to permit the stimulation and/or expansion of T  
CC cells) and inhibiting the development of a cancer in a patient (by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the  
CC polynucleotide, protein or antigen presenting cells that express the  
CC polynucleotide, such that T cells proliferate, administering to the  
CC patient an effective amount of the proliferated T cells, and thus  
CC inhibiting the development of a cancer in the patient. The  
CC polynucleotide, protein and cells are useful in a composition for

CC stimulating an immune response in a patient, and for treating a cancer in  
CC a patient (particularly lung cancer). The oligonucleotide is useful for  
CC determining the presence of a cancer in a patient. The protein and  
CC oligonucleotides are useful in pharmaceutical compositions, e.g.  
CC vaccines. The polynucleotide is also useful as a probe or primer for  
CC nucleic acid hybridisation, and in the design and preparation of  
CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
CC proteins in tumour cells. An amplified portion of the polynucleotide is  
CC useful for isolating a full-length gene from a suitable library.  
CC The present sequence encodes a fusion protein of human RAL2 with  
CC the protein product of a cDNA (full length, extended or partial)  
CC isolated from a library derived from lung tumour/cancer cells.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from the USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=20020197669.

XX  
SQ Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 other;

## Alignment Scores:

Pred. No.: 7.23e-58 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-684-215A-18 (1-128) x ACA12098 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGCAGGATTCGCCATTCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTCCACCGTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCCTCGGCTTGGTGTTCACAAACGCAACGCGCAGGATCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTGATCACCAGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCAACTCGCCACCGGATGGGGAGCGCTTAACGGGCATATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGCAGCGGTACAGGAAACGTG 381  
QY 121 ThrLeuAlaGluGlyProAla 128  
DB 382 ACATGGCCGAGGGACCCCGGCC 405

## RESULT 7

ACA03284

ID ACA03284 standard; DNA; 822 BP.

XX ACA03284;

AC ACA03284;

XX 22-MAY-2003 (first entry)

DT Lung cancer therapy and diagnosis associated DNA #5.

DE Lung cancer; cytostatic; vaccine; gene therapy; cancer;

XX gene; ds.

XX Homo sapiens.

OS US2002172952-A1.

XX

PN

XX 21-NOV-2002.  
PD 10-JUL-2001; 2001US-0902941.  
XX 30-JUN-1999; 99US-0346492.  
PF 15-OCT-1999; 99US-0419356.  
XX 17-DEC-1999; 99US-0466867.  
PR 30-DEC-1999; 99US-0476300.  
PR 06-MAR-2000; 2000US-0519642.  
PR 22-MAR-2000; 2000US-0533077.  
PR 10-APR-2000; 2000US-0546259.  
PR 27-APR-2000; 2000US-0560406.  
PR 05-JUN-2000; 2000US-0589184.  
PR 11-JUL-2000; 2000US-0614124.  
PR 29-AUG-2000; 2000US-0651563.  
PR 08-SEP-2000; 2000US-0658824.  
PR 26-SEP-2000; 2000US-0671325.  
PR 06-OCT-2000; 2000US-0677419.  
PR 30-OCT-2000; 2000US-0702705.  
PR 13-DEC-2000; 2000US-0736457.  
PR 03-MAY-2001; 2001US-0849626.  
XX (CORI-) CORIXA CORP.  
PA Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;  
PI Carter D, Fanger GR, Vedwick TS, Bangur CS, McNabb A;  
PI WPI; 2003-328427/31;  
XX New polynucleotide, useful for preparing a composition for treating or  
PT inhibiting development of cancer, e.g. lung cancer -  
PT Example 8; SEQID NO 1862; 82pp; English.  
PS The invention describes an isolated polynucleotide comprising one of 32  
XX sequences, complement or degenerate variants of them. The polynucleotide  
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
CC for treating or inhibiting development of cancer, e.g. lung cancer.  
CC This sequence represents a polynucleotide associated with the  
CC compositions and methods for the therapy and diagnosis of lung cancer.  
XX Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 other;  
SQ Alignment Scores:  
Pred. No.: 7.23e-58 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0  
US-09-684-215A-18 (1-128) x ACA03284 (1-822)  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGCAGGATTCGCCATTCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTTCACCGTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCCTCGGCTTGGTGTTCACAAACGCAACGCGCAGGATCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTGATCACCAGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCAACTCGCCACCGGATGGGGAGCGCTTAACGGGCATATCCC 321

Qy 101 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGACGCTACAGGAACGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 8

ID ABK27798 standard; cDNA; 894 BP.  
 XX AC ABK27798;  
 XX DT 09-APR-2002 (first entry)  
 XX DE Human colon cancer expressed sequence tag, Seq ID no 236.  
 XX KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;  
 KW expressed sequence tag.  
 XX OS Homo sapiens.  
 XX PN WO200196390-A2.  
 XX PD 20-DEC-2001.  
 XX PF 08-JUN-2001; 2001WO-US18577.  
 XX PR 09-JUN-2000; 2000US-210821P.  
 PR 18-DEC-2000; 2000US-256571P.  
 PR 10-MAY-2001; 2001US-290240P.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;  
 DR WPI; 2002-139708/18.  
 DR P-PSDB; AAU81532.

Novel isolated polynucleotide encoding a polypeptide comprising a portion of colon tumour protein, useful for detection, diagnosis and therapy of human colon cancer -  
 Claim 1; Page 215-216; 220pp; English.  
 The invention relates to an isolated polynucleotide (I) encoding a polypeptide (II) comprising at least a portion of a colon tumour protein. (I), (II) and antibody (III) to (II) are useful for determining the presence of a cancer in a patient. (I), (II) or antigen presenting cells expressing (I) is useful for stimulating and/or expanding T cells specific for a tumour protein, by contacting T cells with (I), (II) or antigen-presenting cells that express (I), under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. (I), (II), or antigen presenting cells that express (II) are useful for treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated from a patient with (I), (II) or antigen presenting cells that express (II), such that T cells proliferate, and administering to the patient an effective amount of the proliferated T cells, thus inhibiting the development of a cancer in the patient. (I) or (II) is useful in vaccines and pharmaceutical compositions for prevention and treatment of colon malignancies and for the diagnosis and monitoring of such cancers. (I), (II) or (III) is useful for detection, diagnosis and/or therapy of human colon cancer. (I) is useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of ribozyme molecules for inhibiting expression of (II) in tumour cells. ABK27364-ABK27807 represent novel human colon cancer coding sequences and primers of the invention.

Sequence 894 BP; 207 A; 269 C; 237 G; 181 T; 0 other;

Alignment Scores:

Pred. No.: 7.99e-58 Length: 894  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-684-215A-18 (1-128) x ABK27798 (1-894)  
 Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCCGCAATCCCATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCCTACC 141  
 Qy 41 AlaphLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCTCGCTGGGTGGTGTGTCGACACACGCAACGCGCAGCAGTCCAAACGCGTG 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGAGCGCTCGCGCGCAAGTCTCGGATCTCCACCGCGGACGTGATCACCAGCGTC 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
 Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCATCATCCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGCAGCGTACAGGAACGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGACCCCGGCC 405  
 RESULT 9  
 ABQ92443  
 ID ABQ92443 standard; DNA; 900 BP.  
 XX AC ABQ92443;  
 XX DT 07-OCT-2002 (first entry)  
 XX DE Human lung cancer associated DNA sequence SEQ ID NO:353.  
 KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine;  
 KW gene; ds.  
 XX OS Homo sapiens.  
 XX PN WO200247534-A2.  
 XX PD 20-JUN-2002.  
 XX PF 30-NOV-2001; 2001WO-US47576.  
 XX PR 12-DEC-2000; 2000US-0735705.  
 PR 07-MAY-2001; 2001US-0850716.  
 PR 28-JUN-2001; 2001US-0897778.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;  
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;  
 DR WPI; 2002-583465/62.  
 DR P-PSDB; ABP61920.  
 XX PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by the polynucleotides, useful in pharmaceutical compositions such as

PT vaccines and as markers to indicate the presence of lung cancer -  
 PS Claim 8; Page 340-341; 381pp; English.  
 XX  
 CC The present invention describes isolated human lung carcinoma  
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
 CC activity, and can be used in gene therapy and in vaccines. Compositions  
 CC comprising (I) or (II) can be used for stimulating an immune response in  
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of  
 CC (I) can be used for detecting the presence of a cancer in a patient, by  
 CC obtaining a biological sample from the patient, contacting the  
 CC biological sample with the oligonucleotide, detecting in the sample, an  
 CC amount of polynucleotide that hybridises to the oligonucleotide and  
 CC comparing the amount of polynucleotide that hybridises to the  
 CC oligonucleotide to a predetermined cut-off value, and determining the  
 CC presence of a cancer in the patient. (I) and (II) are useful in  
 CC pharmaceutical compositions, e.g. vaccines. (I) is useful as a marker to  
 CC indicate the presence or absence of a cancer such as lung cancer.  
 CC ABQ92145 to ABQ92486 and ABP61866 to ABP61992 represent sequences used  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,06e-58 Length: 900  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABQ92443 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCTCGGCTTGGTGTTCGACAAACACGACGCGGAGTCCCAACGCGGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCCGGAGGCTCCGGCGCAAGTCTCGGCATCTCCACGCGGACGTGATCACCGGGTCC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120  
 Db 322 GTGAGCTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGTACAGGGAACGTG 381  
 QY 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 10

ABL49257  
 ID ABL49257 standard; cDNA; 900 BP.

XX

AC ABL49257;

XX

DT 01-MAY-2002 (first entry)

XX

DE Ral2/C-terminal portion of L763p fusion polynucleotide SEQ ID NO:353.  
 XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW Immune response; ss.

XX  
 OS Homo sapiens.  
 XX WO200200174-A2.  
 PN  
 XX 03-JAN-2002.  
 PD  
 XX 28-JUN-2001; 2001WO-US21065.  
 XX  
 PR 28-JUN-2000; 2000US-0606421.  
 PR 02-AUG-2000; 2000US-0630940.  
 PR 21-AUG-2000; 2000US-0643597.  
 PR 15-SEP-2000; 2000US-0662786.  
 PR 09-OCT-2000; 2000US-0685696.  
 PR 12-DEC-2000; 2000US-0735705.  
 PR 07-MAY-2001; 2001US-0850716.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Wang A, SKeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;  
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
 XX  
 DR WPI: 2002-090513/12.  
 DR P-PSDB; ABB75000.  
 XX  
 PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response -  
 XX  
 PS Claim 1; Page 333-334; 374pp; English.  
 XX  
 CC The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies,  
 CC fusion proteins, T cell populations, or antigen presenting cells that  
 CC express the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL49959 to ABL49300 and ABB74946 to  
 CC ABB75070 represent sequences used in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 other;

Alignment Scores:

Pred. No.: 8,06e-58 Length: 900  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABL49257 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCTCGGCTTGGTGTTCGACAAACACGACGCGGAGTCCCAACGCGGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCCGGAGGCTCCGGCGCAAGTCTCGGCATCTCCACGCGGACGTGATCACCGGGTCC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321

```
OY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
D 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCAGCGGTACAGGGAACGTG 381
OY 121 ThrLeuAlaGluGlyProProAla 128
D 382 ACATTGGCCGAGGAGGCCCGGCC 405

RESULT 11
AAS64141
ID AAS64141 standard; cDNA; 915 BP.
XX
AC AAS64141;
XX
XX 29-JAN-2002 (first entry)
XX
DE Human /M. tuberculosis Ra12 fusion protein RA12-P775P-ORF3 cDNA.
XX
KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Microbacterium tuberculosis.
OS Synthetic.
XX
XX WO200173032-A2.
XX
PD 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US09919.
XX
XX 27-MAR-2000; 2000US-0536857.
XX
PR 09-MAY-2000; 2000US-0568100.
XX
PR 12-MAY-2000; 2000US-0570737.
XX
PR 13-JUN-2000; 2000US-0593793.
XX
PR 27-JUN-2000; 2000US-0605783.
XX
PR 10-AUG-2000; 2000US-0636215.
XX
PR 29-AUG-2000; 2000US-0651236.
XX
PR 06-SEP-2000; 2000US-0657279.
XX
PR 02-OCT-2000; 2000US-0679426.
XX
PR 10-OCT-2000; 2000US-0685166.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
XX WPI; 2001-639232/73.
XX
DR P-PSDB; AAU69902.
XX
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer -
XX
XX Example 17; Page 536-537; 579pp; English.
XX
XX The invention relates to isolated prostate-specific
CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of
CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC the antigen-presenting cells are useful for stimulating and/or expanding
CC T cells specific for a tumour protein, and for inhibiting the development
CC of cancer especially prostate cancer. Compositions comprising the
CC polynucleotide and/or polypeptide are useful for stimulating an immune
CC response, and for treating cancer. The oligonucleotide is useful for
CC detecting cancer. The present sequence is a prostate specific
CC polynucleotide of the invention.
XX
SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;
XX
```

Alignment Scores:

Pred. No.: 8.22e-58 Length: 915

```
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-684-215A-18 (1-128) x AAS64141 (1-915)

OY 1 ThrAlaAlaSerAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
D 22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCCGATCCGATC 81
OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
D 82 GGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
OY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
D 142 GCCTTCCTCGGCTTGGGTGTGTGCGACAACAACGGCAACGGCGACGAGTCCACGGCGTG 201
OY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
D 202 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTGATCACCGCGGTC 261
OY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
D 262 GACGGCGCTCCGATCAACTCGGCACCGCGATGCGGACGCGCTTAACGGGCGATCATCCC 321
OY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
D 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCAGCGGTAAACGGGCGATCATCCC 381
OY 121 ThrLeuAlaGluGlyProProAla 128
D 382 ACATTGGCCGAGGAGGCCCGGCC 405

RESULT 12
AAS64141
ID AAS64141 standard; cDNA; 915 BP.
XX
XX AAS64141;
XX
XX 04-OCT-2001 (first entry)
XX
XX Ra12-P775P-ORF3 construct cDNA sequence.
DE Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis; ss.
XX
XX Homo sapiens.
OS
XX WO200151633-A2.
XX
XX 19-JUL-2001.
XX
XX 16-JAN-2001; 2001WO-US01574.
XX
XX 14-JAN-2000; 2000US-0483672.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Mesagher MJ;
XX
XX WPI; 2001-425673/45.
XX
XX New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
XX Claim 8; Page 497-498; 543pp; English.
XX
```

CC The present invention describes polynucleotide sequences (I) which encode  
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AH93357 to AH93944 and AA01115 to  
 CC AA01318 represent polynucleotide and amino acid sequences used in the  
 CC exemplification of the present invention.

XX  
 SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;

## Alignment Scores:

Pred. No.: 8,22e-58 Length: 915  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-684-215A-18 (1-128) x AAH93905 (1-915)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 22 ACGGCGGCTCCGATAACTCCAGCTGTCCAGGTGGCAGGGATTCCGCATTCGGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 DB 82 GGGCAGGCGATCGCATCGCGGCCAGATCAAGCTCCACCGTTTCATATCGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 DB 142 GCCTTCCTCGGCTTGGGTGTTCGACACACAGCGGCGGCGGATCCCAACGCGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 DB 202 GTCCGGAGGCGTCCGGCGCAAGTCTCGGCATCTCCACGGCGAGTGATCACCGCGTCC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 DB 262 GACGGCGTCCGATCAACTCGGCCACCGGATGGCGGAGCGGCTTAACGGGCATCATCC 321  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 DB 322 GGTGACGTCATCTCGTGACCTGGCAACCAAGTCCGGCGGCGGCGGTACAGGGAGCGTG 381  
 QY 121 ThrLeuAlaGluGlyProProAla 128  
 DB 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 13

ABL95512  
 ID ABL95512 standard; cDNA; 915 BP.

XX AC ABL95512;

XX DT 19-JUL-2002 (first entry)

XX DE Ra12- P775P-ORF3 construct cDNA sequence SEQ ID NO 834.

XX KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;

XX KW gene therapy; gene; ss.

XX OS Chimeric - Mycobacterium tuberculosis.

XX OS Chimeric - Homo sapiens.

XX PN US2002022248-A1.

XX

PD 21-FEB-2002.  
 XX  
 PF 12-JAN-2001; 2001US-0759143.  
 XX  
 PR 25-FEB-1997; 97US-0806099.  
 PR 01-AUG-1997; 97US-0904804.  
 PR 09-FEB-1998; 98US-0020956.  
 PR 25-FEB-1998; 98US-0030607.  
 PR 14-JUL-1998; 98US-0115453.  
 PR 23-SEP-1998; 98US-0159812.  
 PR 15-JAN-1999; 99US-0232149.  
 PR 09-APR-1999; 99US-0288946.  
 PR 13-JUL-1999; 99US-0352616.  
 PR 12-NOV-1999; 99US-0439313.  
 PR 18-NOV-1999; 99US-0443686.  
 PR 14-JAN-2000; 2000US-0483672.  
 PR 27-MAR-2000; 2000US-0536857.  
 PR 09-MAY-2000; 2000US-0568100.  
 PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.  
 PR 27-JUN-2000; 2000US-0605783.  
 PR 10-AUG-2000; 2000US-0636215.  
 PR 29-AUG-2000; 2000US-0651236.  
 PR 06-SEP-2000; 2000US-0657279.  
 PR 02-OCT-2000; 2000US-0679426.  
 PR 10-OCT-2000; 2000US-0685166.

XX (XUJJ/) XU J.  
 PA (DILL/) DILLON D C.  
 PA (MITC/) MITCHAM J L.  
 PA (HARL/) HARLOCKER S L.  
 PA (JIAN/) JIANG Y.  
 PA (KALO/) KALOS M D.  
 PA (FANG/) FANGER G R.  
 PA (RETT/) RETTER M W.  
 PA (STOL/) STOLK J A.  
 PA (DAYC/) DAY C H.  
 PA (VEDV/) VEDVICK T S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKEL/) SKEIKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 XX WPI; 2002-255649/30.

XX New prostate-specific polynucleotides for diagnosing and treating  
 PT diseases, in particular prostate cancer, and as markers for the  
 PT progression of cancer

PS Example 17; SEQ ID NO 834; 87pp; English.

XX The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a cDNA  
 CC described in the invention.

SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;

## Alignment Scores:

Pred. No.: 8,22e-58 Length: 915  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABL95512 (1-915)



Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGCGTCCGATACCTCCAGCTGCTCCAGGGTGGCAGGATTCGCATTCGGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGCGATGCGATGCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
 Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCCTCGCTTGGTGTGTTCGACACAAACGCGGACGCGGATCCACCGGGTG 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGCTCGGCGGCAAGTCTCGACAAACGCGGACGCGGATCCACCGGGTG 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGCTCCGATCACTCCGCGCACCGGATGCGGACGCGCTTAACGGGATCATCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGGTACAGGGAAGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGGAGGACCCCGGGC 405

RESULT 14  
 ACAS9949  
 ID ACAS9949 standard; cdna; 915 BP.  
 XX AC  
 AC ACAS9949;  
 DT 10-JUN-2003 (first entry)  
 XX DE  
 DE Prostate cancer therapy associated cdna #650.  
 XX KW  
 KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA;  
 KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 KW PSMA; gene; ss.  
 XX OS  
 OS Homo sapiens.  
 XX PN  
 PN US2002192763-A1.  
 XX PD  
 PD 19-DEC-2002.  
 XX PF  
 PF 29-JUN-2001; 2001US-0895793.  
 XX PR  
 PR 17-APR-2000; 2000US-157455P.  
 PR 04-OCT-2000; 2000US-0679272.  
 PR 28-MAR-2001; 2001US-0822827.  
 XX PA  
 PA (XUJ/) XU J.  
 PA (DILL/) DILLON D C.  
 PA (MITC/) MITCHAM J L.  
 PA (HARL/) HARLOCKER S L.  
 PA (JIAN/) JIANG Y.  
 PA (KALO/) KALOS M D.  
 PA (FANG/) FANGER G R.  
 PA (RETT/) RETTER M W.  
 PA (STOL/) STOLK J A.  
 PA (DAYC/) DAY C H.  
 PA (VEDV/) VEDVICK T S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.

PA (HURA/) HURAL J.  
 PA (MCNE/) MCNEILL P D.  
 PA (HOUG/) HOUGHTON R L.  
 PA (DBAS/) Y DE BASSOLS C V.  
 PA (FOYT/) FOY T M.  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
 PI Mcneill PD, Houghton RL, Y De Bassols CV, Foy TM;  
 XX WPI: 2003-352711/33.  
 XX New fusion protein comprising prostate-specific polypeptides, or its  
 PT immunogenic portions, useful for diagnosing, preventing and/or treating  
 PT cancer, particularly prostate cancer  
 XX Example 17; SEQ ID NO 834; 85pp; English.  
 PS The invention describes a fusion protein comprising at least one amino  
 CC acid sequence of immunogenic portions of any of the 3 sequences not  
 CC defined in the specification, or sequences having at least 70 or 90 %  
 CC sequence identity to any one of the 35 sequences defined in the USPTO  
 CC web site, which is encoded by any of the 4 nucleotide sequences not  
 CC defined in the specification. The fusion protein, composition and  
 CC methods are useful for diagnosing, preventing and/or treating cancer,  
 CC particularly prostate cancer. The proteins are useful as markers to  
 CC indicate the presence or absence of cancer. This sequence  
 CC represents a prostate cancer therapy associated cdna.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=US20020192763.  
 XX SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8.22e-58 Length: 915  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0  
 US-09-684-215A-18 (1-128) x ACA59949 (1-915)  
 Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGCGTCCGATACCTCCAGCTGCTCCAGGGTGGCAGGATTCGCATTCGGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGCGATGCGATGCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
 Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCCTCGCTTGGTGTGTTCGACAAACGCGGACGCGGATCCACCGGGTG 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGCTCGGCGGCAAGTCTCGACAAACGCGGACGCGGATCCACCGGGTG 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGCTCCGATCACTCCGCGCACCGGATGCGGACGCGCTTAACGGGATCATCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGGTACAGGGAAGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGGAGGACCCCGGGC 405

## RESULT 15

ABK39768  
ID ABK39768 standard; cDNA; 945 BP.XX  
AC ABK39768;

XX DT 21-MAY-2002 (first entry)

XX DE DNA encoding lung tumour protein p801p ORF4 and Ra12 fusion protein.

XX KW Lung tumour; cancer; T cell; immune response stimulator;  
KW Cytostatic; gene; ss.XX OS Homo sapiens.  
OS Synthetic.

XX PN WO200204514-A2.

XX PD 17-JAN-2002.

XX PF 10-JUL-2001; 2001WO-US22058.

XX PR 11-JUL-2000; 2000US-0614124.

XX PR 29-AUG-2000; 2000US-0651563.

XX PR 08-SEP-2000; 2000US-0658824.

XX PR 26-SEP-2000; 2000US-0671325.

XX PR 06-OCT-2000; 2000US-0677419.

XX PR 30-OCT-2000; 2000US-0702705.

XX PR 13-DEC-2000; 2000US-0736457.

XX PR 03-MAY-2001; 2001US-0849626.

XX PA (CORI-) CORIXA CORP.

XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;

PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;

PI Mcnabb A, Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;

XX WPI; 2002-164634/21.

DR P-PSDB; AAU85586.

XX Novel polynucleotide encoding a lung tumour polypeptide useful for  
PT stimulating and/or expanding T cells specific for a tumour protein

XX Example 8; SEQ ID No 1861; 223pp; English.

XX The invention describes an isolated polynucleotide and polypeptide  
CC useful for stimulating and/or expanding T cells specific for a tumour  
CC protein for determining the presence of a cancer in a patient. A  
CC composition containing the polynucleotide and/or polypeptide is useful  
CC for treating a lung cancer in a patient. The polypeptide is useful for  
CC removing tumour cells from a biological sample. The polynucleotide is  
CC also useful as probe or primer to detect the level of mRNA encoding a  
CC tumour protein. This sequence encodes a lung tumour associated protein  
CC or protein fragment, described in the method of the invention.CC Note: the sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 945 BP; 150 A; 318 C; 360 G; 117 T; 0 other;

## Alignment Scores:

Pred. No.:	8.54e-58	Length:	945
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-684-215A-18 (1-128) x ABK39768 (1-945)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

Db	22	ACGGCCGGTCCGATAACTTCCAGGTGTCAGGGTGGCAGGGATTCCCGCATTCGGATC 81
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db	82	GGGCAGGGGATGGCCATCGCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141
Qy	41	AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db	142	GCCTTCTCTGGCTTGGGTGTTGTCGACAAACGCAACGGGGCAGAGTCCCAACGCGTG 201
Qy	61	ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db	202	GTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGGACGTGATCACCGCGGTC 261
Qy	81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db	262	GACGGCGTCCGATCAACTCGGCCACCGCGGATGGCGGACGCTTAACGGGGCATCATCCC 321
Qy	101	GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db	322	GGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGGCGGCACGCGTACAGGGAACGTG 381
Qy	121	ThrLeuAlaGlyGlyProProAla 128
Db	382	ACATTGGGCGGAGGGACCCCGCGCC 405

Search completed: September 5, 2003, 09:04:19

Job time : 250.83 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:28:59 ; Search time 68.4138 Seconds  
(without alignments)  
825.814 Million cell updates/sec

Title: US-09-684-215A-18

Perfect score: 653

Sequence: 1 TAASDNFQLSQGGQGAIP1.....QTKSGTRTGNVTIAEGPPA 128

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_593/app\_query.fasta\_1.853  
-DB-Issued\_Patents\_NA -QWTF-fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -LOCAL=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215\_8CGN\_1\_145 @runat\_05092003\_072203\_593 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	653	100.0	822	4	US-09-736-457-1862
2	653	100.0	900	4	Sequence 1862, App
3	653	100.0	900	4	Sequence 353, App
4	653	100.0	945	4	US-09-606-421B-353
5	653	100.0	1012	4	Sequence 1861, App
6	653	100.0	1012	4	US-09-736-457-1861
7	653	100.0	1464	4	Sequence 351, App
8	653	100.0	1464	4	US-09-606-421B-351
9	653	100.0	1557	4	Sequence 348, App
10	653	100.0	1557	4	US-09-620-412C-348
11	653	100.0	1578	4	Sequence 332, App
12	653	100.0	1578	4	US-09-556-877-188
					Sequence 188, App

13	653	100.0	1578	4	US-09-598-419-188
14	653	100.0	1752	4	Sequence 188, App
15	653	100.0	1752	4	US-09-620-412C-352
16	653	100.0	1758	4	Sequence 352, App
17	653	100.0	1758	4	US-09-598-419-352
18	653	100.0	1860	4	Sequence 336, App
19	653	100.0	1860	4	US-09-620-412C-336
20	653	100.0	1896	4	Sequence 336, App
21	653	100.0	1896	4	US-09-598-419-336
22	653	100.0	1941	4	Sequence 308, App
23	653	100.0	1941	4	US-09-620-412C-308
24	653	100.0	1965	4	Sequence 324, App
25	653	100.0	1965	4	US-09-598-419-324
26	653	100.0	2052	4	Sequence 316, App
27	653	100.0	2052	4	US-09-620-412C-316
28	653	100.0	2076	4	Sequence 316, App
29	653	100.0	2076	4	US-09-598-419-316
30	653	100.0	2103	4	Sequence 340, App
31	653	100.0	2103	4	US-09-620-412C-340
32	653	100.0	2148	4	Sequence 340, App
33	653	100.0	2148	4	US-09-598-419-340
34	653	100.0	2148	4	Sequence 356, App
35	653	100.0	2148	4	US-09-620-412C-356
36	632	96.8	2287	4	Sequence 312, App
37	632	96.8	4403765	3	US-09-598-419-312
38	632	96.8	4411529	3	US-09-620-412C-312
39	627	96.0	447	4	Sequence 312, App
40	627	96.0	447	4	US-09-598-419-312
41	627	96.0	447	4	Sequence 344, App
42	627	96.0	447	4	US-09-620-412C-344
43	627	96.0	1872	3	Sequence 320, App
44	627	96.0	1872	3	US-09-598-419-320
45	627	96.0	1872	4	Sequence 328, App
					US-09-620-412C-328
					Sequence 320, App
					US-09-598-419-320
					Sequence 328, App
					US-09-620-412C-328
					Sequence 328, App
					US-09-598-419-328
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 1, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 17, Appli
					Sequence 17, Appli
					Sequence 17, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-736-457-1862

Sequence 1862, Application US/09736457

Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick

APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun

APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1862

LENGTH: 822

TYPE: DNA

ORGANISM: Homo sapiens

US-09-736-457-1862

Alignment Scores:

Pred. No.: 1.9e-64

Score: 653.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 4

Length: 822

Matches: 128

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0





```
|||||
Db      427 ACATTGGCGGAGGACCCCGGCC 450

RESULT 6
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-351

Alignment Scores:
Pred. No.:      2,49e-64      Length:      1012
Score:          653.00      Matches:      128
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            4      Gaps:      0

US-09-684-215A-18 (1-128) x US-09-606-421B-351 (1-1012)
QY      1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
|||||
Db      67 ACGGGCGGCTCCGATAAATTCAGCTGTCCAGGGTGGCAGGGATTCCGATTCGGATC 126
|||||
QY      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
|||||
Db      127 GGGCAGCGCATCGGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 186
|||||
QY      41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
|||||
Db      187 GCCTTCCTCGGCTGGGTGTTCGACAAACACGACGAGCGGCGCAGATCCCAACGCGTG 246
|||||
QY      61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
|||||
Db      247 GTCGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGGGTC 306
|||||
QY      81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
|||||
Db      307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCGATCATCC 366
|||||
QY      101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db      367 GTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACCGGTACAGGGAACGTG 426
|||||
QY      121 ThrLeuAlaGluGlyProProAla 128
|||||
Db      427 ACATTGGCGGAGGACCCCGGCC 450

RESULT 7
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Alignment Scores:
Pred. No.:      4,06e-64      Length:      1464
Score:          653.00      Matches:      128
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            4      Gaps:      0

US-09-684-215A-18 (1-128) x US-09-620-412C-348 (1-1464)
QY      1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
|||||
Db      22 ACGGGCGGCTCCGATAAATTCAGCTGTCCAGGGTGGCAGGGATTCCGATTCGGATC 81
|||||
QY      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
|||||
Db      82 GGGCAGCGCATCGGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
|||||
QY      41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
|||||
Db      142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGACGAGCGGCGCAGATCCCAACGCGTG 201
|||||
QY      61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
|||||
Db      202 GTCGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGGGTC 261
|||||
QY      81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
|||||
Db      262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCGATCATCC 321
|||||
QY      101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db      322 GGTGACGCTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACCGGTACAGGGAACGTG 381
|||||
QY      121 ThrLeuAlaGluGlyProProAla 128
|||||
Db      382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 8
US-09-598-419-348
; Sequence 348, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-348

Alignment Scores:
```

Pred. No.: 4,06e-64 Length: 1464  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-348 (1-1464)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
DB 22 AGCGCGCGCTCCGATTAACCTCCAGGTGGCAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
DB 82 GGCAGCGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
DB 142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCAACGGCGGCGACGATCCAAACCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
DB 202 GTCCGGAGCGCTCCGCGCGCAAGTCTCCGATCTCCACCGGCGACGTGATCACCAGCGGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
DB 262 GACGGCGCTCCGATCAACTCCGCGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
DB 322 GGTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
DB 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 9

US-09-620-412C-332  
; Sequence 332, Application US/09620412C  
; Patent No. 6448234

; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 332

; LENGTH: 1557

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-620-412C-332

Alignment Scores:  
Pred. No.: 4,4e-64 Length: 1557  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-620-412C-332 (1-1557)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
DB 22 AGCGCGCGCTCCGATTAACCTCCAGGTGGCAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
DB 82 GGCAGCGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141

QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
DB 142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCGCAACGGCGGCGACGATCCAAACCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
DB 202 GTCCGGAGCGCTCCGCGCGCAAGTCTCCGATCTCCACCGGCGGACGATCCACCGGGTGC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
DB 262 GACGGCGCTCCGATCAACTCCGCGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
DB 322 GGTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
DB 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 10

US-09-598-419-332

; Sequence 332, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C6

; CURRENT APPLICATION NUMBER: US/09/598,419

; CURRENT FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 357

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 332

; LENGTH: 1557

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-598-419-332

Alignment Scores:  
Pred. No.: 4,4e-64 Length: 1557  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-332 (1-1557)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
DB 22 AGCGCGCGCTCCGATTAACCTCCAGGTGGCAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
DB 82 GGCAGCGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
DB 142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCGCAACGGCGGCGACGATCCAAACCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
DB 202 GTCCGGAGCGCTCCGCGCGCAAGTCTCCGATCTCCACCGGCGGACGATCCACCGGGTGC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
DB 262 GACGGCGCTCCGATCAACTCCGCGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||





Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-188 (1-1578)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCGCAATCCGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141

Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCAGCGATCCACCGCGGTC 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCGCGGCAAGTCTCGGCATCTCCACCGCGAGCGTATCACCAGCGGTC 261

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGACGGCGCTTAACGGGCATCATCCC 321

Qy 101 GlyAspValIleSerValThrTrrPlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTACGCTCATCTCGGTGACCTGGCAAAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProAla 128  
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 14

US-09-620-412C-352  
; Sequence 352, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620.412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 352  
; LENGTH: 1752  
; TYPE: DNA  
; ORGANISM: Chlamydia trachomatis  
US-09-620-412C-352

Alignment Scores:  
Pred. No.: 5.14e-64 Length: 1752  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-620-412C-352 (1-1752)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCGCAATCCGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141

Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCAGCGATCCACCGCGGTC 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCGCGGCAAGTCTCGGCATCTCCACCGCGAGCGTATCACCAGCGGTC 261

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGACGGCGCTTAACGGGCATCATCCC 321

Qy 101 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141

Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCAGCGATCCACCGCGGTC 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCGCGGCAAGTCTCGGCATCTCCACCGCGAGCGTATCACCAGCGGTC 261

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGACGGCGCTTAACGGGCATCATCCC 321

Qy 101 GlyAspValIleSerValThrTrrPlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTACGCTCATCTCGGTGACCTGGCAAAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProAla 128  
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 15

US-09-598-419-352  
; Sequence 352, Application US/09598419  
; Patent No. 6565856  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Scholler, John  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C6  
; CURRENT APPLICATION NUMBER: US/09/598.419  
; CURRENT FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 357  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 352  
; LENGTH: 1752  
; TYPE: DNA  
; ORGANISM: Chlamydia trachomatis  
US-09-598-419-352

Alignment Scores:  
Pred. No.: 5.14e-64 Length: 1752  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-352 (1-1752)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCGCAATCCGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141

Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCAGCGATCCACCGCGGTC 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCGCGGCAAGTCTCGGCATCTCCACCGCGAGCGTATCACCAGCGGTC 261

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGACGGCGCTTAACGGGCATCATCCC 321

Qy 101 GlyAspValIleSerValThrTrrPlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTACGCTCATCTCGGTGACCTGGCAAAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProAla 128  
|||||  
Db 382 ACATTGGCCGAGGGACCCCGGCC 405

Search completed: September 5, 2003, 12:12:10  
Job time : 75.4138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:52:39 ; Search time 2285.46 Seconds  
(without alignments)  
2291.192 Million cell updates/sec

Title: US-09-684-215A-18  
Perfect score: 653  
Sequence: 1 TAASDNFQLSGGGQFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framer\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US09684215/runat\_05092003\_072202\_559/app\_query.fasta.1.853  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215@cgn.1.1.3962 -runat\_05092003\_072202\_559 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: en\_ba.\*  
16: en\_fun.\*  
17: en\_hum.\*  
18: en\_in.\*  
19: en\_mu.\*  
20: en\_om.\*  
21: en\_or.\*  
22: en\_ov.\*  
23: en\_pat.\*  
24: en\_ph.\*  
25: en\_pl.\*  
26: en\_ro.\*  
27: en\_sts.\*  
28: en\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	653	100.0	675	6	AR261272 Sequence
2	653	100.0	675	6	AX201049 Sequence
3	653	100.0	675	6	AX267848 Sequence
4	653	100.0	822	6	AR277645 Sequence
5	653	100.0	822	6	AX369152 Sequence
6	653	100.0	894	6	AX351489 Sequence
7	653	100.0	900	6	AR220690 Sequence
8	653	100.0	900	6	AX365960 Sequence
9	653	100.0	915	6	AR261281 Sequence
10	653	100.0	915	6	AX201061 Sequence
11	653	100.0	915	6	AX267860 Sequence
12	653	100.0	945	6	AR277644 Sequence
13	653	100.0	945	6	AX369151 Sequence
14	653	100.0	1012	6	AR220689 Sequence
15	653	100.0	1012	6	AX365958 Sequence
16	653	100.0	1155	6	AX369165 Sequence
17	653	100.0	1203	6	AX201078 Sequence
18	653	100.0	1203	6	AX267877 Sequence
19	653	100.0	1464	6	AR229410 Sequence
20	653	100.0	1464	6	AX156105 Sequence
21	653	100.0	1464	6	AX361955 Sequence
22	653	100.0	1557	6	AR229398 Sequence
23	653	100.0	1557	6	AX156089 Sequence
24	653	100.0	1557	6	AX361939 Sequence
25	653	100.0	1578	6	AR229328 Sequence
26	653	100.0	1578	6	AX155945 Sequence
27	653	100.0	1578	6	AX361795 Sequence
28	653	100.0	1590	6	AX316986 Sequence
29	653	100.0	1752	6	AR229413 Sequence
30	653	100.0	1752	6	AX156109 Sequence
31	653	100.0	1752	6	AX361959 Sequence
32	653	100.0	1758	6	AR229401 Sequence
33	653	100.0	1758	6	AX156093 Sequence
34	653	100.0	1758	6	AX361943 Sequence
35	653	100.0	1860	6	AR229380 Sequence
36	653	100.0	1860	6	AX156065 Sequence
37	653	100.0	1860	6	AX361915 Sequence
38	653	100.0	1860	6	AR229392 Sequence
39	653	100.0	1896	6	AX156081 Sequence
40	653	100.0	1896	6	AX361931 Sequence
41	653	100.0	1941	6	AR229386 Sequence
42	653	100.0	1941	6	AX156073 Sequence
43	653	100.0	1941	6	AX361923 Sequence
44	653	100.0	1965	6	AR229404 Sequence
45	653	100.0	1965	6	AX156097 Sequence

ALIGNMENTS

RESULT 1

AR261272 LOCUS 675 bp DNA linear PAT 29-JAN-2003  
SEQUENCE 822 from patent US 6321716.  
ACCESSION AR261272  
VERSION AR261272.1 GI:28072035  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 675)  
AUTHORS Mashiki,Z. and Harada,J.  
TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 822 27-NOV-2001;  
FEATURES Location/Qualifiers  
source 1..675  
BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.28e-44 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-18 (1-128) x AR261272 (1-675)  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGGCGCTCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCATGGCGATCGCGGCCAGATCAAGCTTCCACCCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCTCGACAAACACGACGCGGCGAGTCAACGCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGCGAGTGATCACCGCGTTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGTCCGATCAACTCGCCACCGCGATGGCGGCGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCGGAGGACCCCGGCC 405  
RESULT 2  
AX201049 LOCUS 675 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 679 from Patent WO0151633.  
ACCESSION AX201049  
VERSION AX201049.1 GI:15390857  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W., Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.  
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

cancer  
Patent: WO 0151633-A 679 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..675  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.28e-44 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-18 (1-128) x AX201049 (1-675)  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGGCGCTCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCATGGCGATCGCGGCCAGATCAAGCTTCCACCCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCTCGACAAACACGACGCGGCGAGTCAACGCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGCGAGTGATCACCGGTTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGTCCGATCAACTCGCCACCGCGATGGCGGCGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCGGAGGACCCCGGCC 405  
RESULT 3  
AX267848 LOCUS 675 bp DNA linear PAT 26-OCT-2001  
DEFINITION Sequence 822 from Patent WO0173032.  
ACCESSION AX267848  
VERSION AX267848.1 GI:16516494  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H., Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T. and Henderson,R.A.  
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer  
JOURNAL Patent: WO 0173032-A 822 04-OCT-2001;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..675  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

BASE COUNT		162 a		197 c		190 g		126 t	
ORIGIN									
Alignment Scores:									
Pred. No.:	2	28e-44	Length:	675					
Score:	653.00	Matches:	128						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	6	Gaps:	0						
US-09-684-215A-18 (1-128) x AX267848 (1-675)									
Qy	1	ThrAlaAlaSerAspAsnPhleGlnGlyGlnGlyGlnGlyPheAlaIleProIle	20						
Db	22	ACGGCCCGCTCCGATAACTTCCAGCTGTCACAGGTGGCAGGATTCCGATCCGATC	81						
Qy	21	GlyClnAlaMetAlaIleAlaGlyClnIleLysLeuProThrValHisIleGlyProThr	40						
Db	82	GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTATATCGGGCTACC	141						
Qy	41	AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal	60						
Db	142	GCCTTCTCCGCTTGGGTGGTTCGACAAACAGGCAACGGCGCAGTCCACGCGTG	201						
Qy	61	ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal	80						
Db	202	GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGCAGTGCATCCCGCGTC	261						
Qy	81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro	100						
Db	262	GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCATCATCC	321						
Qy	101	GlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArgThrGlyAsnVal	120						
Db	322	GGTGACGTCACTCGGTGACCTGCCAACCAAGTCGGCGCGCACCGCTACAGGGAACGTG	381						
Qy	121	ThrLeuAlaGluGlyProProAla 128							
Db	382	ACATTGGCCGAGGACCCCGGCC 405							
RESULT 4									
AR277645									
LOCUS	AR277645	822 bp	DNA	linear	PAT 10-APR-2003				
DEFINITION	Sequence 1862 from patent US 6509448.								
ACCESSION	AR277645								
VERSION	AR277645.1 GI:29711294								
KEYWORDS	Unknown.								
SOURCE	Unknown.								
ORGANISM	Unclassified.								
REFERENCE	1 (bases 1 to 822)								
AUTHORS	Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mannion, J., Fan, L. and Wang, A.								
TITLE	Compositions and methods for the therapy and diagnosis of lung cancer								
JOURNAL	Patent: US 6509448-A 1862 21-JAN-2003;								
FEATURES	Location/Qualifiers								
source	1..822								
BASE COUNT	146 a	281 c	270 g	125 t					
ORIGIN									
Alignment Scores:									
Pred. No.:	2	79e-44	Length:	822					
Score:	653.00	Matches:	128						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	6	Gaps:	0						
US-09-684-215A-18 (1-128) x AR277645 (1-822)									

QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
Db 142 GCCTTCTCGGTGGTGTCTGCACACACAGCGCAACGGCGCAGGATCCCAACCGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGGACGTATCACCAGCGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
|||||  
Db 262 GAGCGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
Db 382 ACATTGGCGAGGAGGCCCGCGCC 405  
RESULT 6  
AX351489  
LOCUS AX351489 894 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 236 from Patent WO0196390.  
ACCESSION AX351489  
VERSION AX351489.1 GI:18616835  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PATENT: WO 0196390-A 236 20-DEC-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source  
1. .894  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 207 a 269 c 237 g 181 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.05e-44 Length: 894  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-18 (1-128) x AX351489 (1-894)  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGCTCCGATAAATTCACAGCTGTCCAGGTGGCGCAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
Db 82 GGGCAGCGCATCGGATCAACTTCACAGCTGTCCAGGTGGCGCAGGATTCGCCATTCGGATC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
Db 142 GCCTTCTCGGTGGTGTCTGCACACACAGCGCAACGGCGCAGGATTCGCCATTCGGATC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGGACGTATCACCAGCGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
|||||  
Db 262 GAGCGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
Db 382 ACATTGGCGAGGAGGCCCGCGCC 405  
RESULT 8  
AX365960

Db 262 GACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
Db 382 ACATTGGCGAGGAGGCCCGCGCC 405  
RESULT 7  
AR220690  
LOCUS AR220690 900 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 353 from patent US 6426072.  
ACCESSION AR220690  
VERSION AR220690.1 GI:23327471  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PATENT: US 6426072-A 353 30-JUL-2002;  
FEATURES  
source  
1. .900  
/organism="unknown"  
BASE COUNT 219 a 239 c 246 g 196 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.07e-44 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-18 (1-128) x AR220690 (1-900)  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGCTCCGATAAATTCACAGCTGTCCAGGTGGCGCAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
Db 82 GGGCAGCGCATCGGATCAACTTCACAGCTGTCCAGGTGGCGCAGGATTCGCCATTCGGATC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
Db 142 GCCTTCTCGGTGGTGTCTGCACACACAGCGCAACGGCGCAGGATTCGCCATTCGGATC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGGACGTATCACCAGCGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
|||||  
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
Db 382 ACATTGGCGAGGAGGCCCGCGCC 405  
RESULT 8  
AX365960

LOCUS AX365960 900 bp DNA linear PAT 15-FEB-2002  
DEFINITION Sequence 353 from Patent WO200174.  
ACCESSION AX365960  
VERSION AX365960.1 GI:18697458  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A.,  
McNeill, P.D., Fanger, N., Retter, M.W., Marnerakis, M., Fanger, G.R.,  
Wedwick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.  
TITLE Compositions and methods for the therapy and diagnosis of lung  
cancer  
JOURNAL Patent: WO 0200174-A 353 03-JAN-2002;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..900  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 219 a 239 c 246 g 196 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,07e-44 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX365960 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGCAGCGCATGGCGATCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAGCGATGCGGCGCATCATCC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 142 GCCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAGCGATGCGGCGCATCATCC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTGCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGCGACGTGATCACC GGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCCAGCCACCGCGATGGCGAGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
Db 382 ACATTGGCGAGGAGGCCCGCGGCC 405

RESULT 9  
AR261281

LOCUS AR261281 915 bp DNA linear PAT 29-JAN-2003  
DEFINITION Sequence 834 from patent US 6321716.  
ACCESSION AR261281  
VERSION AR261281.1 GI:28072044  
KEYWORDS Unknown.  
SOURCE  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 915)  
AUTHORS Mashiki, Z. and Harada, J.  
TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 834 27-NOV-2001;  
FEATURES  
source Location/Qualifiers  
1..915  
/organism="unknown"

BASE COUNT 167 a 280 c 284 g 184 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.12e-44 Length: 915  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AR261281 (1-915)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGCAGCGCATGGCGATCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAGCGATGCGGCGCATCATCC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTGCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGCGACGTGATCACC GGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCCAGCCACCGCGATGGCGAGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
Db 382 ACATTGGCGAGGAGGCCCGCGGCC 405

RESULT 10  
AX201061

LOCUS AX201061 915 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 691 from Patent WO0151633.  
ACCESSION AX201061  
VERSION AX201061.1 GI:15390868  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W.,  
Stolk, J.A., Skeiky, Y.A., Wang, A. and Meagher, M.J.  
TITLE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0151633-A 691 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..915  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 167 a 280 c 284 g 184 t

```
ORIGIN
Alignment Scores:
Pred. No.: 3.12e-44 Length: 915
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX201061 (1-915)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGGCTCCGATAACTTCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGCTTGGGTGTTGCGACAACACGCGCACGCGGCGAGTCCACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 CTCGGGAGCGTCCGGCGCAAGTCTCGGCATCTCCACCGCGAGTGCATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGCTTAACGGCGATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGAGGACCGCGGCC 405

RESULT 11
AX267860
LOCUS AX267860 915 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 834 from Patent WO0173032.
ACCESSION AX267860
VERSION AX267860.1 GI:16516503
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H., Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T. and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0173032-A 834 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..915
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 167 a 280 c 284 g 184 t
ORIGIN

Alignment Scores:
Pred. No.: 3.12e-44 Length: 915
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX201061 (1-915)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGGCTCCGATAACTTCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGCTTGGGTGTTGCGACAACACGCGCACGCGGCGAGTCCACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 CTCGGGAGCGTCCGGCGCAAGTCTCGGCATCTCCACCGCGAGTGCATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGCTTAACGGCGATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGAGGACCGCGGCC 405

RESULT 12
AX277644
LOCUS AR277644 945 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1861 from patent US 6509448.
ACCESSION AR277644
VERSION AR277644.1 GI:29711293
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 945)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 1861 21-JAN-2003;
FEATURES
Location/Qualifiers
1..945
/organism="unknown"

BASE COUNT 150 a 318 c 360 g 117 t
ORIGIN

Alignment Scores:
Pred. No.: 3.23e-44 Length: 945
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AR277644 (1-945)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGGCTCCGATAACTTCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
```



```

Qy 41 AlaphetLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGCTGGGTGGTGTTCGACAAACAGCGGCAACGGCGAGTCCAAACGGGT 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGGACGATGATCACCAGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGATCACTCGGCACCGGATGGCGGCGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACGCTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCCGAGGAGCCCGCGGC 405
RESULT 13
LOCUS AX369151 945 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1861 from Patent WO0204514.
ACCESSION AX369151
VERSION AX369151.1 GI:18857169
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
Marnerakis,M., Carter,D., Fanger,G.R., Vedwick,T.S., Bangur,C.S.,
McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 1861 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source 1. 945
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 318 c 360 g 117 t
ORIGIN
Alignment Scores:
Pred. No.: 3.23e-44 Length: 945
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-18 (1-128) x AX369151 (1-945)
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCGCGCTCCGATACTTCCAGCTGTCCAGGTTGGCGGCGGATTCGCCATTCGCGATC 81
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGCATGGCGATCGGGCGGCGAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 141
Qy 41 AlaphetLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGCTGGGTGGTGTTCGACAAACAGCGGCAACGGCGAGTCCAAACGGGT 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGGACGATGATCACCAGCGGTC 261

```

```

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGATCACTCGGCACCGGATGGCGGCGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACGCTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCCGAGGAGCCCGCGGC 405
RESULT 14
LOCUS AR220689 1012 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 351 from patent US 6426072.
ACCESSION AR220689
VERSION AR220689.1 GI:23327470
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6426072-A 351 30-JUL-2002;
FEATURES
source Location/Qualifiers
BASE COUNT 235 a 289 c 289 g 199 t
ORIGIN
Alignment Scores:
Pred. No.: 3.47e-44 Length: 1012
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-18 (1-128) x AR220689 (1-1012)
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 67 ACGCGCGCTCCGATACTTCCAGCTGTCCAGGTTGGCGGCGGATTCGCCATTCGCGATC 126
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 127 GGGCAGCGCATGGCGATCGGGCGGCGAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 186
Qy 41 AlaphetLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 187 GCCTTCCTCGCTGGGTGGTGTTCGACAAACAGCGGCAACGGCGAGTCCAAACGGGTG 246
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 247 GTCGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGGCGAGTATCACCAGCGTC 306
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 307 GACGGCGCTCGATCACTCGGCACCGGATGGCGGCGGCTTAACGGGCATCATCCC 366
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 367 GGTGACGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACGCTACAGGGAACGTG 426
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 427 ACATTGGCCGAGGAGCCCGCGGC 450
RESULT 15

```

AX365958  
 LOCUS AX365958 1012 bp DNA linear PAT 15-FEB-2002  
 DEFINITION Sequence 351 from Patent WO200174.  
 ACCESSION AX365958  
 VERSION AX365958.1 GI:18697457  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 REFERENCE Wang, T., Wang, A., Skeiky, Y. A., Li, S. X., Kalos, M. D., Henderson, R. A.,  
 McNeill, P. D., Fanger, N., Retter, M. W., Marnerakis, M., Fanger, G. R.,  
 Vedvick, T. S., Carter, D., Watanabe, Y. and Peckham, D. W.  
 TITLE Compositions and methods for the therapy and diagnosis of lung  
 cancer  
 JOURNAL Patent: WO 0200174-A 351 03-JAN-2002;  
 CORIXA CORPORATION (US)  
 FEATURES  
 source  
 1..1012  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 235 a 289 c 289 g 199 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3 47e-44 Length: 1012  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-684-215A-18 (1-128) x AX365958 (1-1012)  
 QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 67 ACGGCGGCTCCGATAAATCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 126  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40  
 DB 127 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 186  
 QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 DB 187 GCCTTCCCTCGGCTTGGGTGTTCGACAACAACGCGCACGGCGCAGTCCACGCGTG 246  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 DB 247 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTGCATCACCGCGTC 306  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 DB 307 GACGGCGCTCCGATCACTCGCCACCGGATGGCGGAGCGGCTTAACGGGATCATCCC 366  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 DB 367 GGTGACGTATCTCGTACCTTGGCAACCAAGTCGGCGGCGACGCGTACAGGGAACGTG 426  
 QY 121 ThrLeuAlaGluGlyProAla 128  
 DB 427 ACATTGGCGGAGGAGACCCCGGCC 450

Search completed: September 5, 2003, 10:31:42  
 Job time : 2287.8 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:27:24 ; Search time 2211.31 Seconds  
(without alignments)  
1406.846 Million cell updates/sec

Title: US-09-684-215A-18  
Perfect score: 653  
Sequence: 1 TAASDNFOLSQGQGFAPFI.....OTKSGGTRGNTVLAEGPPA 128

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlip  
-Q=/cgn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072202\_572/app\_query.fasta\_1.853  
-DB=EST -OPM=fastp -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMF=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -CGN\_1\_1\_4575\_@runat\_05092003\_072202\_572 -NCPU=6 -ICPU=3  
-NO\_MMAB -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	% Query	Length	DB	ID	Description
1	145.5	22.3	289	29	U82114	U82114	orde
2	125.5	19.2	1438	28	BH770798	BH770798	LMGtag54
3	123	18.8	959	29	BZ549048	pacs1-60	BZ549048
4	122.5	18.8	726	14	CB679186	OSJNEF02F	CB679186
5	119.5	17.3	590	14	CB925602	ABAL_22.F	CB925602
6	114	17.5	758	13	BQ514888	EST622303	BQ514888
7	106	16.2	603	28	AZ934428	BJ_Ba000	AZ934428
8	104	15.9	765	9	AJ558965	AJ558965	AJ558965
9	100.5	15.4	1033	29	BZ561390	pacs1-164	BZ561390
10	100	15.3	947	29	BZ549047	pacs1-60	BZ549047
11	99	15.2	726	28	AQ989479	RfC00025	AQ989479
12	98	15.0	1093	13	BU557763	AGENCOURT	BU557763
13	96.5	14.8	790	29	BZ577987	msb2_5667	BZ577987
14	96	14.7	546	9	AW285510	LGI_241_E	AW285510
15	96	14.7	551	9	AW285527	LGI_241_G	AW285527
16	95.5	14.6	677	12	BI378928	BFLG1_000	BI378928
17	95	14.5	543	12	BI721127	1031054B1	BI721127
18	95	14.5	673	12	BI956121	HVSMEM002	BI956121
19	94.5	14.5	1127	29	BZ560689	pacs2-164	BZ560689
20	93.5	14.3	885	29	BZ567978	pacs2-184	BZ567978
21	92	14.1	703	14	CD383492	PTMM08663	CD383492
22	91.5	14.0	600	10	BG143271	la95h02.y	BG143271
23	91	13.9	701	14	CD204316	HS1_7_H09	CD204316
24	91	13.9	1098	12	BU553475	AGENCOURT	BU553475
25	90.5	13.9	728	12	BJ285991	BJ285991	BJ285991
26	89	13.6	509	10	BE195909	HVSMEM009	BE195909
27	88	13.5	655	14	CB921057	VVD048E11	CB921057
28	88	13.5	999	12	BM051288	603634270	BM051288
29	87.5	13.4	600	29	CC345457	OCQAF707H	CC345457
30	87.5	13.4	655	9	AL883551	AL883551	AL883551
31	87.5	13.4	719	28	AZ933900	BJ_Ba000	AZ933900
32	87.5	13.4	898	10	BF309796	601892449	BF309796
33	87	13.3	510	13	BQ977384	QHI24M07	BQ977384
34	87	13.3	642	12	BJ283688	BJ283688	BJ283688
35	87	13.3	643	12	BJ254998	BJ254998	BJ254998
36	87	13.3	649	13	BQ801199	WHE2811.E	BQ801199
37	87	13.3	656	13	BU016293	OHE12L01	BU016293
38	87	13.3	799	12	BM943807	UI-M-EH0P	BM943807
39	87	13.3	1373	13	BQ890329	AGENCOURT	BQ890329
40	86.5	13.2	491	10	BE333846	us28a04.y	BE333846
41	86.5	13.2	494	10	BF722271	mab16h10	BF722271
42	86.5	13.2	518	12	BM569886	lh99c11.y	BM569886
43	86.5	13.2	547	9	AI652699	AT652699	AT652699
44	86.5	13.2	555	12	BI439305	ic62807.y	BI439305
45	86.5	13.2	555	12	BI441548	ic55908.y	BI441548

ALIGNMENTS

RESULT 1  
LOCUS U82114  
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.  
ACCESSION U82114  
VERSION U82114.1  
KEYWORDS GI:3647212  
SOURCE GSS.  
ORGANISM Mycobacterium leprae  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacterium.  
REFERENCE 1 (bases 1 to 289)

U82114 289 bp DNA linear GSS 16-FEB-2001  
U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.

**AUTHORS** Silbaq,F.S., Cho,S.N., Cole,S.N., Cole,S.T. and Brennan,P.J.  
**TITLE** Characterization of a 34-kilodalton protein of *Mycobacterium leprae* that is isologous to the immunodominant 34-kilodalton antigen of *Mycobacterium paratuberculosis*  
**JOURNAL** Infect. Immun. 66 (11), 5576-5579 (1998)  
**MEDLINE** 99003183  
**PUBMED** 9784577  
**COMMENT** Contact: Silbaq FS  
 Microbiology  
 Colorado State University  
 Fort Collins, CO 80523, USA  
 Elgimeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae*. Mol. Microbiol. 7 (2), 197-206 (1993)  
 Class: unknown.

**FEATURES** source  
 1..289  
 Location/Qualifiers  
 /organism="Mycobacterium leprae"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1769"  
 /clone="cosmid L-373; contig 64"  
 /clone.lib="ordered cosmid library"  
 BASE COUNT 62 a 83 c 93 g 51 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7,94e-06 Length: 289  
 Score: 145.50 Matches: 35  
 Percent Similarity: 58.82% Conservative: 15  
 Best Local Similarity: 41.18% Mismatches: 34  
 Query Match: 22.28% Indels: 1  
 DB: 29 Gaps: 1

US-09-684-215A-18 (1-128) x U82114 (1-289)

QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsn---GlyAlaArgValGlnArg 59  
 ||| |||||: ||| : ||| |||||: |||  
 Db 7 GCGTGGTGGTGTACAGTGGCCACCACCAAGCGCCGCGGCGCAAGATTATGAC 66  
 ||||| : ||| : ||| |||||: |||  
 QY 60 ValValGlySerAlaProAlaAlaSerLeuGlyLeuSerThrGlyAspValIleThrAla 79  
 ||||| : ||| : ||| |||||: |||  
 Db 67 GTCGTGGCGGTGGTGGCGCGCGGATCCCGCGTCCCAAGGGGTCTCTACTAAG 126  
 ||||| : ||| : ||| |||||: |||  
 QY 80 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 99  
 ||||| : ||| : ||| |||||: |||  
 Db 127 GTCGACGACGCGCTGATCAGTAGCGCGCGGTTGGTGGTGGTGGTGGTGGTGGTGGT 186  
 ||||| : ||| : ||| |||||: |||  
 QY 100 ProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsn 119  
 ||||| : ||| : ||| |||||: |||  
 Db 187 CCCGTGACAAGTGTCTGACCTATCAGGATCAGTCTGGTAGCAGTCCGACGTTTCCAG 246  
 ||||| : ||| : ||| |||||: |||  
 QY 120 ValThrLeuAlaGlu 124  
 ||||| : ||| : ||| |||||: |||  
 Db 247 GTCACACTCGGCAAG 261  
 ||||| : ||| : ||| |||||: |||

**RESULT 2**  
 BH770798/c 1438 bp DNA linear GSS 01-MAY-2002  
**LOCUS** LLMGtag541 MG1363 Random Sequence Tag Library Lactococcus lactis  
**DEFINITION** subsp. cremoris genomic, genomic survey sequence.  
**ACCESSION** BH770798  
**VERSION** BH770798.1 GI:20373755  
**KEYWORDS** GSS.  
**SOURCE** Lactococcus lactis subsp. cremoris  
**ORGANISM** Lactococcus lactis subsp. cremoris  
 Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;  
 Lactococcus.

**REFERENCE** 1 (bases 1 to 1438)  
**AUTHORS** Bolotin,A., Ehrlich,S.D. and Sorokin,A.  
**TITLE** Studies of genomes of dairy bacteria *Lactococcus lactis*  
**JOURNAL** Sci. Aliments, (2002) In press  
**COMMENT** Contact: Sorokin A  
 Genetique Microbienne

## INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France  
 Tel: 33 1 34 65 25 16  
 Fax: 33 1 34 65 25 21  
 Email: sorokine@jouy.inra.fr  
 best homologue in strain IL1403 is htrA (95%)  
 Class: shotgun  
 High quality sequence start: 30  
 High quality sequence stop: 1408.

## FEATURES source

Location/Qualifiers  
 1..1438  
 /organism="Lactococcus lactis subsp. cremoris"  
 /mol\_type="genomic DNA"  
 /strain="MG1363"  
 /db\_xref="taxon:1359"  
 /clone.lib="MG1363 Random Sequence Tag Library"  
 /note="Vector: pSGMU2; Site:1: SmaI; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."  
 BASE COUNT 438 a 282 c 260 g 458 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.00723 Length: 1438  
 Score: 125.50 Matches: 40  
 Percent Similarity: 43.28% Conservative: 18  
 Best Local Similarity: 29.85% Mismatches: 45  
 Query Match: 19.22% Indels: 31  
 DB: 28 Gaps: 5

US-09-684-215A-18 (1-128) x BH770798 (1-1438)

QY 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla----- 27  
 ||| |||||: ||||| : ||| |||||: |||  
 Db 479 GGTTAGGTTTGGTATCCCATCATGATGCTGTAACATCATTAATAAATTTCAAACT 420  
 ||||| : ||| : ||| |||||: |||  
 QY 28 ---GlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly 46  
 |||||: ||| : ||| |||||: |||  
 Db 419 GATGTAAGATTTCACGTCCTGCC-----TTAGGTATTTCGT 384  
 ||||| : ||| : ||| |||||: |||  
 QY 47 ValValAsp-----AsnAsnGlyAsn----- 53  
 :||: ||||| : |||||: |||  
 Db 383 ATGGTTGACCTTCTCAATTATCAACAANTAGTCTCTCACTGAAATTTACCTAGCAGC 324  
 ||||| : ||| : ||| |||||: |||  
 QY 54 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGly 70  
 ||||| : ||| : ||| |||||: |||  
 Db 323 GTAACGTGTGGAGTGGTGTCTCTCTCTCAAGCGGTCTTCTGCTGCCACAGCTGGT 264  
 ||||| : ||| : ||| |||||: |||  
 QY 71 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 90  
 :||: |||||: ||||| : ||| |||||: |||  
 Db 263 CTGAAGCTGGCGATGTGATAACGAAGTGGGAGATACCGCGTACTTCAACACAGAC 204  
 ||||| : ||| : ||| |||||: |||  
 QY 91 MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThr 110  
 :||: ||||| : ||||| : ||| |||||: |||  
 Db 203 TTACAAGTGTCTTTTACTCTACACAAATATTAAATGATCTGTGAAGTCACTACTAC 147  
 ||||| : ||| : ||| |||||: |||  
 QY 111 LysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGlu 124  
 :||: ||||| : ||||| : ||| |||||: |||  
 Db 146 CGTGATGTAATAATCAGCCACAGCAATGTTCAAACTCTCTATAA 105  
 ||||| : ||| : ||| |||||: |||

**RESULT 3**  
 BZ549048/c 959 bp DNA linear GSS 17-DEC-2002  
**LOCUS** pacsl-60\_1644.s2 pacsl-60 Pseudomonas aeruginosa genomic clone  
**DEFINITION** pacsl-60\_1644, genomic survey sequence.  
**ACCESSION** BZ549048  
**VERSION** BZ549048.1 GI:27152629  
**KEYWORDS** GSS.  
**SOURCE** Pseudomonas aeruginosa  
**ORGANISM** Pseudomonas aeruginosa  
 Bacteria: Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.

**REFERENCE** 1 (bases 1 to 959)

**AUTHORS** Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,  
Burns,J.L., Kaul,R. and Olsen,M.V.

**TITLE** Whole-Genome-Sequence variation among multiple isolates of  
*Pseudomonas aeruginosa* library

**JOURNAL** J. Bacteriol., (2002) In press

**COMMENT** Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 20622216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

**FEATURES**

source	1..959	Location/Qualifiers
	/organism="Pseudomonas aeruginosa"	
	/mol_type="genomic DNA"	
	/strain="1-60"	
	/db_xref="taxon:287"	
	/clone_lib="pacsl-60"	
	/clone_lib="pacsl-60"	
	/note="clinical isolate 1-60 whole genomic shotgun library."	

BASE COUNT	157 a	314 c	291 g	195 t	2 others
ORIGIN					

Alignment Scores:

Pred. No.:	0.00791	Length:	959
Score:	123.00	Matches:	41
Percent Similarity:	44.88%	Conservative:	16
Best local Similarity:	32.28%	Mismatches:	46
Query Match:	18.84%	Indels:	24
DB:	29	Gaps:	4

US-09-684-215A-18 (1-128) x BZ549048 (1-959)

```

QY      4 SerAspAsnPhedInLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db     378 TCGCAGATCTTACCCGGTTCCGGCGGTTCATGGCGCTGTCTTCGCATTCCGATCGAT 319
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
QY     22 GlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla 41
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db    318 GTCCGCGCTGAACGTGCCCGACCAGTTCGAAAG-----AAAGCGGCAAGTCACT 271
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
QY     42 PheLeuGlyLeuGlyValValalaspAsnAsnGlyAsn----- 53
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db    270 CGCGCGTGCGTGGCGGTGGTGCACGAGAAGTGAACAAGGATCTCGCGGAGTCTCATCGGC 211
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
QY     54 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 68
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db    210 CTCGACAAAGCCGTCCGCGCGTGGTGGCGCACTGTGGGAAGAGCGCCGCGGCCCAAG 151
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
QY     69 LeuGlyIleSerThrGlyAspValIleIleThraValAlaAspGlyAlaProIleAsnSerAla 88
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db    150 GGTTGGCTGCAGTGGCGCATGTGATCATTAAGCTTGAAGCCAGTCGATCAACGAGTCC 91
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
QY     89 ThrAlaMetAlaaspAlaLeuasnGlyHishisProGlyaspValIleSerValThrTrp 108
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db     90 GCCGACTCGCCGCACTGTGGTGGCAACATGAAGCCGGCGCAAGATCAACCTG----- 37
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
QY    109 GlnThrLysSerGlyGlyThr 115
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db     36 -----CACGGGGGATCC 25
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||

```

**RESULT 4**  
CB679186/c

**LOCUS** CB679186.1

**DEFINITION** OSJNEF02F19 r OSJNEF Oryza sativa (japonica cultivar-group) CDNA clone OSJNEF02F19 3', mRNA sequence.

**ACCESSION** CB679186

**VERSION** CB679186.1

**KEYWORDS** GI:29682911

**SOURCE** EST.  
Oryza sativa (japonica cultivar-group)  
EST.

```

ORGANISM
Oryza sativa (Japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzoae; Oryza.

REFERENCE
1 (bases 1 to 726)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 02 row: F column: 19
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
source          1..726
                /organism="Oryza sativa (Japonica cultivar-group)"
                /mol_type="mRNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="OSJNEF03F19"
                /tissue_type="Leaf"
                /dev_stage="3 week"
                /lab_host="DH10B"
                /clone_lib="OSJNEf"
                /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                XhoI; Uninfected Control"

BASE COUNT      221 a   181 c   151 g   173 t
ORIGIN

Alignment Scores:
Pred. No.:       0.00629      Length:       726
Score:           122.50
Percent Similarity: 42.86%    Matches:      41
Best Local Similarity: 30.83% Mismatches:     16
Query Match:      18.76%     Indels:       27
DB:               14         Gaps:        4

US-09-684-215A-18 (1-128) x CB679186 (1-726)

Qy      10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
      ::: ||| ||||| ||||| ||||| |::: ||||| |||
Db      579 ACATCTGCTGGTGTGGTTTGCCATCCCGTCATCAACTGTTCTGAAAATAGCTCCTCAG 520
      :::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuValValAsp 49
      ::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      519 TTA-----ATTCA GTTTGGAAGAAGTTCGCCGTGCTGGCTGTAATGTGCAGTTT 472
      :::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      50 -----ASnAsnClyAsnGlyAlaArgValGlnArgVal 60
      :::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      471 GCTCCAGATCCAATCGCATATCATGCTTAATTTGCGCACTGGATCTATAAATTGACGGTT 412
      :::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      61 ValGlySerAlaProAlaAlaSerLeuGly----- 70
      ||| ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      411 CTTGGGGGCGAGTGCTGCAGCCAAANGCTGGTCTTGTTCCTACCAGTAGGGGTTTTGCTGGT 352
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      351 ACCATTGTTCTTGGTGATGTCATTTGTTGCCGTGGACGTAAACCTATCAAAGGCAAAATCT 292
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      90 AlaMetAlaSpAlaLeuAsnGlyHisHisProClyAspValIleSerValThrTrpGln 109
      :::::: ||||| :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      291 GATCTATFCAGGGGTTTCGGATGCTTATFGGCGTTGGAGACAAGGTGAGCTTGACAATCCAA 232
      :::::: ||||| :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY      110 ThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
Db      231 AGA-----GGCGCTGAATCTCTGGAGGTAACCTTG 202

RESULT 5
CB925602
LOCUS   CB925602
DEFINITION
ABAI_22_F06.b1_A012 Abscissic acid-treated seedlings Sorghum bicolor
cDNA clone ABAI_22_F06_A012 3', mRNA sequence.
ACCESSION
CB925602
VERSION
CB925602.1 GI:30161873
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 590)
Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S., Klein
,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
Eastman,A. and Pratt,L.H.
An EST database from Sorghum: ABAI-treated seedlings
Unpublished
Other_ESTs: ABAI_22_F06.g1_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug3 (CGACCTGCAGCTCGAGCACCA)
POLYA=yes.

FEATURES
Location/Qualifiers
1..590
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="IS3620C"
/db_xref="taxon:4558"
/clone="ABAI_22_F06_A012"
/lab_host="PH10B-n1 phage-resistant E. coli"
/clone_lib="Abscissic acid-treated seedlings"
/notes="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. After 12 days, medium was
supplemented with 1 mM abscissic acid (ABA), while leaves
were misted with a solution of 1 mM ABA. Roots and leaves
were harvested after 3, 6, 12, and 24 hr and material from
all time points was combined prior to RNA isolation.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5'-prime
DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG
). XhoI excises the cDNA insert."
BASE COUNT 136 a 137 c .156 g 161 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00995 Length: 590
Score: 119.50 Matches: 38
Percent Similarity: 43.41% Conservative: 18
Best Local Similarity: 29.46% Mismatches: 40
Query Match: 18.30% Indels: 33
DB: 14 Gaps: 5

US-09-684-215A-18 (1-128) x CB925602 (1-590)

```

```

QY      15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThr 34
Db      6 GCCTTGTCTATTCATCATCACTGCTACTTAAATCGCTCCTCAGTTA-----53

RESULT 6
BQ514888
LOCUS   BQ514888
DEFINITION
ESTG22303 Generation of a set of potato cDNA clones for microarray
analyses mixed potato tissues Solanum tuberosum cDNA clone STMIO81
3' end, mRNA sequence.
ACCESSION
BQ514888
VERSION
BQ514888.1 GI:21373757
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished
Other_ESTs: EST622302
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7.

FEATURES
Location/Qualifiers
1..758
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="kennebec or Binjete"
/db_xref="taxon:4113"
/clone="STMIO81"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of stolons, leaves, leaflets,
infestans-treated libraries of stolons, leaflets,
axillary buds of stem explants, petioles, germinating eyes

```

BASE COUNT 215 a 181 c 128 g 234 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 0.051 Length: 758  
Score: 114.00 Matches: 34  
Percent Similarity: 42.61% Conservative: 15  
Best Local Similarity: 29.57% Mismatches: 42  
Query Match: 17.46% Indels: 24  
DB: 13 Gaps: 3

US-09-684-215A-18 (1-128) x BQ514888 (1-758)

Qy 10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29  
Db 592 ACATCAGCAGGTGTGGATTGCAATCCCTTCCTCACTGTGTGAGAGTGTGCCCCAG 533  
Qy 30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49  
Db 532 TTG-----ATCCAATCTGGAAAGTTCTTCGTGCTGTTTGAATATTGAAATC 485  
Qy 50 -----AsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 484 GCTCCAGACCTGATGTCACCAACCACTTAATGTTGCAATGAGCACTGGTTTCTGGTA 425  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGly-----TCTACTACCGGGTTTTCAGGA 365  
Db 424 CCTGGAATAGTCTGCGAGCGAAAGCGGACTTCTTCTACTACCGGGTTTTCAGGA 365  
Qy 71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89  
Db 364 AATATAGTGTGCGGATATTATTGAAGCAGTGTGATGACAAACCTGTTAGGAGTAAAGCA 305  
Qy 90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 104  
Db 304 GAGTTGTATAAGCCCTGGTAACTATAACATAGTGTGAAGTT 260

RESULT 7  
AZ934428 603 bp DNA linear GSS 24-APR-2001  
LOCUS  
DEFINITION BJ\_Ba0002108r B. japonicum BAC library Bradyrhizobium japonicum genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

ORGANISM  
Bradyrhizobium japonicum  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Bradyrhizobiaceae; Bradyrhizobium.

REFERENCE  
AUTHORS  
TITLE

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Genome Res. 11 (8), 1434-1440 (2001)  
21376150  
11483585  
Contact: Wing RA  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Class: BAC ends  
High quality sequence stop: 553.

FEATURES  
source  
Location/Qualifiers

1. .603  
/organism="Bradyrhizobium japonicum"  
/mol\_type="genomic DNA"  
/strain="USD110"  
/db\_xref="taxon:375"

/lab\_host="E. coli"  
/clone\_lib="B. japonicum BAC library"  
/note="Vector: pIndigo536; Site\_1: HindIII"

BASE COUNT 103 a 201 c 203 g 90 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 0.261 Length: 603  
Score: 106.00 Matches: 35  
Percent Similarity: 41.80% Conservative: 16  
Best Local Similarity: 28.69% Mismatches: 55  
Query Match: 16.23% Indels: 16  
DB: 28 Gaps: 4

US-09-684-215A-18 (1-128) x AZ934428 (1-603)

Qy 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLys--- 31  
Db 245 GGCAATCGCTTCTCGATCCCGGCCAACACCGTGAAGACGGTTGTGCCAGCTCAAGGAC 304  
Qy 32 -----LeuProThrValHisIleGlyProThrAlaPheLeuGly 44  
Db 305 AAGGTTCTGCTACCCCGCTGGATCGCGCTGCAGATTTCAGCGGTGACG----- 355  
Qy 45 LeuGlyValValAspAsnAsnGly-----AsnGlyAlaArgValGlnArgVal 60  
Db 356 TCGGATATCCCGACAGCCTCGGCATGAAGAAGCGCGGCGCTGGTGGCGAGCCG 415  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 416 CAGCGCAACCGTCCGCGCGAAGCGCGCATCGAGTCCCGCGAGTACACCTCGCTCGTC 475  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
Db 476 AACGGCAATCCGTCAGGAGCGCGCGAGCTCGCCCGCACCATCGCGCGATGGCGCC 535  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 536 CGTGGATCGTGAAGCTTAACGTGCTGCACAAG---GGCCAGGACAATGCTGTAACCTC 592  
Qy 121 ThrLeu 122  
Db 593 ACCCTC 598

RESULT 8

AJ558965

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1. .765

/organism="Antirrhinum majus"

/mol\_type="mrna"

/db\_xref="taxon:4151"

/clone="018\_09\_a08"

/tissue\_type="whole plant"

AJ558965 765 bp mRNA linear EST 12-JUN-2003  
AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone  
018\_09\_a08, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

ORGANISM  
Antirrhinum majus (snapdragon)  
Antirrhinum majus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Lamiaceae; Antirrhinaceae; Antirrhineae;  
Antirrhinum.

REFERENCE  
1 (bases 1 to 765)  
Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.  
Antirrhinum EST collection  
Unpublished  
Contact: Schwarz-Sommer Z  
Molekulare Pflanzen-genetik  
MPI fuer Zuechtungs-forschung  
Carl-von-Linne Weg 10, D-50829, Germany.

FEATURES  
source  
Location/Qualifiers

1. .765  
/organism="Antirrhinum majus"  
/mol\_type="mrna"  
/db\_xref="taxon:4151"  
/clone="018\_09\_a08"  
/tissue\_type="whole plant"

4 others

```

Length:      .    1033
Matches:      42
Conservative: 17
Mismatch:     54
Indels:       25
Gaps:         4

```

1-1033)

Qy	4	SerAspAsnProGlnLeuSerGlnGlyGln-----GlyPheAlaIleProIleGly	21
Db	713	TCCACAGTTTTCACCCGGTCCGCCGGGTCAATGCCCTGCTGTTCGGCATTCGCCATCGAT	654
Qy	22	GlnAlaMetAlaIle-AlaGlyGlnIleLys-LeuProThrValHisIleGlyProThrA	41
		:::	
Db	653	GTCGCGCTCAGCGTTCCGCGACCACTTGAAGAAAGCCC-----GGCAAGGTCA	606
Qy	41	laphLeuGlyValLeuGlyValValAspAsnAsnGlyAsn-----	53
		:::	
Db	605	GTCGCGCGTGGCTGGCGCTGTGTATCCAGGAAGTGAACAAGGATCTCCCGAGTCTCCTCG	546
Qy	54	-----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlas	68
Db	545	GCCTCGACAAGCCGTCGCGCGTGTGTGGCGCAACTGGTGGAAAGACGGCCGCGCGCCA	486
Qy	68	erLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSera	88
Db	485	AGGTGGGCTTCAGGTGGCGCATGTATCTCAACCTGAAACGGCCAGTCTCATCAACGAT	426
Qy	88	lathAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal----	106
		:::	
Db	425	CCGCGCGACCTGCCGACCTGTGGGCAACATGAAGCGCGGCGGCAACAGATCAACTGCAGC	366

Qv

365 TGATTGCAAGGGCCAGCGCAAGTCCCTGAGCATGGCGGTAGGCAGCGCTT 316

BZ3490417C  
LOCUS

**DEFINITION** pacsl-60\_1649.sl pacsl-60 Pseudomonas aeruginosa genomic clone  
pacsl-60\_1649.gis pacsl-60 Pseudomonas aeruginosa genomic clone  
**ACCESSION** pacsl-60\_1649, genomic survey sequence.  
**VERSION** BZ549047  
**KEYWORDS** BZ549047.1 GI:27152628 GSS.

**KEYWORDS:**  
**SOURCE:**

**SYNOPSIS**  
Aeruginosa  
Pseudomonas aeruginosa  
Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.  
1 (bases 1 to 947)  
**REFERENCE**  
Spencer, D. H., Raymond, C. K., Smith, E. E., Sims, E. E., Hastings, M. J.

## AUTHORS

**TITLE** Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*

**JOURNAL** J. Bacteriol., (2002) In press

**COMMENT** Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA

FEATURES  
 source  
 1. 6947  
 Location/Qualifiers  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="1-60"  
 Class: shotgun.  
 Email: craymondeu.washington.edu  
 Fax: 2066857244  
 Tel: 2062280904





BZ577987	BZ577987	790 bp	DNA	linear	GSS 17-DEC-2002
LOCUS	msh2_5667.x1 msh	<i>Pseudomonas aeruginosa</i>	genomic clone	msh2_5667,	
DEFINITION	genomic survey sequence.				
ACCESSION	BZ577987				
VERSION	BZ577987.1	GI:27213048			
KEYWORDS	GSS.				
SOURCE	<i>Pseudomonas aeruginosa</i>				
ORGANISM	<i>Pseudomonas aeruginosa</i>				
REFERENCE	Bacteria: Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.				
AUTHORS	1 (bases 1 to 790)				
	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,				
	Burns,J.L., Kaul,R. and Olsen,M.V.				
TITLE	Whole-Genome Sequence variation among multiple isolates of <i>Pseudomonas aeruginosa</i> library				
JOURNAL	J. Bacteriol., (2002) In press				
COMMENT	Contact: Chris K. Raymond				
	Genome Center				
	University of Washington				
	Box 352145, Seattle, WA 98105-2145, USA				
	Tel: 20622216954				
	Fax: 2066857244				
	Email: craymond@u.washington.edu				
	Class: shotgun.				

```

FEATURES             Location/Qualifiers
     source           1..790
                        /organism="Pseudomonas aeruginosa"
                        /mol_type="genomic DNA"
                        /strain="MSH"
                        /db_xref="taxon:287"
                        /clone="msh2_5667"
                        /clone_lib="msh"
                        /notes="Environmental isolate. Whole genomic shotgun library."
BASE COUNT          107 a   243 c    287 g   150 t      3 others
ORIGIN
Alignment Scores:
Pred. No.:         3.58              Length:       790
Score:            96.50              Matches:      37
Percent Similarity: 22.40%            Conservative:  16
Best Local Similarity: 29.60%        Mismatches:   47
Query Match:      14.78%            Indels:      25
DB:               29                Gaps:        6

US-09-684-215A-18 (1-128) x BZ577987 (1-790)

Qy      9 LeuSerGlnGlyGlnGlnGlyPheAlaIleProIleGly-----GlnAla 23
      |||:::||||||| | | | | | | | | | | | | | | | | | | | |
Db     110 CTTGCCGCCTCAGGGCAGCACACCGCCAGCAGGAGGAGGTCTCGTTCGTTCGCC 169

Qy     24 MetAlaIleAlaGlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPhe 42
      ::::::::::::::::::::||| ||| ||| ::::::::::::::
Db     170 GTCCCGCGTCCCAGGGAATCTGNATTGTATCCGGCCGTACTGGTCGGTCAGATCTCCTC 229

Qy     43 LeuGlyLeuGlyValValAspAsnAnsnGlyAlaArgValGlnArgvalValGly 62
      ||| ||| :|:::||| | | | | | | | | | | | | | | | | | |
Db     230 GCCGCGCTTGCC-----CAACCACGAGGCGGTCTGGGTTCGCATCGCGCGCTGGG 280

Qy     63 SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValaspGly 82
      ||| ||| :|:::||| | | | | | | | | | | | | | | | | | |
Db     281 TGTGACCCGGGGGGCGGTAGCGGTGGCGCTTGGGGATCGCCTCGAACGGGTT----- 334

Qy     83 AlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisProGlyAsp 102
      ||| ::::| ||| | | | | | | | | | | | | | | | | | |
Db     335 CGCGTAGTGGCGTGCGTGGCGTGGCGTGGCGGAC-----CGAGGTCAGCCACCCAGTCGAT 388

Qy    103 ValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrClyAsnValThrLeu 122
      ||| :|::: ||| | | | | | | | | | | | | | | | | | |
Db     389 ATTCAACGAATCGTC-----GTCAATGCCGCC 412

Qy    123 AlaGlucglyPropro 127

```

```

Db      413 GCTGAGGTGAACCA 427
|||||
RESULT 14
LOCUS   AW285510                      546 bp mRNA linear EST 19-JUL-2000
DEFINITION LG1_241_E05.g1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
SEQUENCE
ACCESSION AW285510
VERSION   AW285510.1 GI:6675354
KEYWORDS EST.
SOURCE   Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 546)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 126
High quality sequence stop: 546
POLYA=Yes.

FEATURES             Location/Qualifiers
source               1..546
                    /organism="Sorghum bicolor"
                    /mol_type="mRNA"
                    /db_xref="taxon:4558"
                    /clone_lib="Light Grown 1 (LGI)"
                    /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                    seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI
                    ; The library was made from poly-A RNA in the cloning
                    vector lambda ZAP II. Clones to be sequenced were
                    prepared by mass excision."
BASE COUNT          121 a 127 c 149 g 149 t
ORIGIN
Alignment Scores:
Pred. No.:          2.53          Length:          546
Score:              96.00          Matches:         25
Percent Similarity: 47.62%          Conservative:    15
Best Local Similarity: 29.76%          Mismatches:     26
Query Match:        14.70%          Indels:         18
DB:                  9              Gaps:            2

US-09-684-215A-18 (1-128) x AW285510 (1-546)

Qy      50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
|||||
Db      56 ANTGTTCGCAACGGAGCTTTATCTTAAGTACCTGGGGGCGAGTCTGCAGCGCAAGCA 115
|||||
Qy      70 GlyIleSer-----
|||||
Db      116 GGTCTTGCTCCACCGCGAGGGTTTGTCTGGTAATATTCTTCTGGGTGATATCATCGTT 175
|||||
Qy      79 AlavalaspGlyAlaProIleAsnSerAlaThralametalaspAlaLeuAsnGlyHis 98
|||||
Db      176 GCAGTGGACGGCAACCTGTTAAGGCGCAATCTGACCTGTGAGGGTCTCGGATGACTAT 235
|||||
Qy      99 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
|||||
Db      236 GGCGTCGAGATCAGGTGACCTTGACA-----ATCCGCGGAGGC 274
|||||
Qy      119 AsnValThrLeu 122
|||||
Db      275 TCAGAAACCCCTT 286
|||||

RESULT 15
LOCUS   AW285527                      551 bp mRNA linear EST 19-JUL-2000
DEFINITION LG1_241_G05.g1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
SEQUENCE
ACCESSION AW285527
VERSION   AW285527.1 GI:6675371
KEYWORDS EST.
SOURCE   Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 551)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 16
High quality sequence stop: 552
POLYA=Yes.

FEATURES             Location/Qualifiers
source               1..551
                    /organism="Sorghum bicolor"
                    /mol_type="mRNA"
                    /db_xref="taxon:4558"
                    /clone_lib="Light Grown 1 (LGI)"
                    /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                    seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI
                    ; The library was made from poly-A RNA in the cloning
                    vector lambda ZAP II. Clones to be sequenced were
                    prepared by mass excision."
BASE COUNT          124 a 127 c 150 g 150 t
ORIGIN
Alignment Scores:
Pred. No.:          2.56          Length:          551
Score:              96.00          Matches:         25
Percent Similarity: 47.62%          Conservative:    15
Best Local Similarity: 29.76%          Mismatches:     26
Query Match:        14.70%          Indels:         18
DB:                  9              Gaps:            2

US-09-684-215A-18 (1-128) x AW285527 (1-551)

Qy      50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
|||||
Db      61 ANTGTTCGCAACGGAGCTTTATCTTAAGTACCTGGGGGCGAGTCTGCAGCGCAAGCA 120
|||||
Qy      70 GlyIleSer-----
|||||
Db      121 GGTCTTGCTCCACCGCGAGGGTTTGTCTGGTAATATTCTTCTGGGTGATATCATCGTT 180
|||||
Qy      79 AlavalaspGlyAlaProIleAsnSerAlaThralametalaspAlaLeuAsnGlyHis 98
|||||
Db      181 GCAGTGGACGGCAACCTGTTAAGGCGCAATCTGACCTGTGAGGGTCTCGGATGACTAT 240
|||||
Qy      99 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
|||||

```

```

Db      241 GCGTCGGAGATCAGGTGACCTTGACA-----ATCCGGCGGAGGC 279
QY      119 AsnValThrLeu 122
Db      280 TCAGAAACCCCTT 291

```

Search completed: September 5, 2003, 11:54:41  
Job time : 2216.31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:54:40 ; Search time 185.821 Seconds  
(without alignments)  
1584.308 Million cell updates/sec

Title: US-09-684-215A-18

Perfect score: 653

Sequence: 1 TAASDNFQLSQGGQGAIFI.....QTKSGGTRTGNVTLAGPPA 128

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_P/US09684215/runat\_05092003\_072203\_619/app\_query.fasta\_1.853  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -MAP=US09684215 -eGcn1.1.130 -runat\_05092003\_072203\_619  
-NCPU=6 -ICPU=3 -NO\_MMAR -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	653	100.0	675	9	US-09-759-143-822 Sequence 822, App

2	653	100.0	675	9	US-09-780-669-822 Sequence 822, App
3	653	100.0	675	9	US-09-822-827-822 Sequence 822, App
4	653	100.0	675	10	US-09-895-793-822 Sequence 822, App
5	653	100.0	675	10	US-09-895-793-822 Sequence 822, App
6	653	100.0	675	12	US-09-845-814-822 Sequence 822, App
7	653	100.0	675	12	US-10-144-678A-822 Sequence 822, App
8	653	100.0	822	10	US-09-736-457-1862 Sequence 1862, App
9	653	100.0	822	10	US-09-902-941-1862 Sequence 1862, App
10	653	100.0	822	10	US-09-849-626-1862 Sequence 1862, App
11	653	100.0	822	14	US-10-017-754-1862 Sequence 1862, App
12	653	100.0	900	9	US-09-735-705-353 Sequence 353, App
13	653	100.0	900	10	US-09-850-716A-353 Sequence 353, App
14	653	100.0	900	10	US-09-897-778-353 Sequence 353, App
15	653	100.0	900	12	US-10-117-982-353 Sequence 353, App
16	653	100.0	915	9	US-09-759-143-834 Sequence 834, App
17	653	100.0	915	9	US-09-780-669-834 Sequence 834, App
18	653	100.0	915	9	US-09-822-827-834 Sequence 834, App
19	653	100.0	915	10	US-09-895-793-834 Sequence 834, App
20	653	100.0	915	10	US-09-895-814-834 Sequence 834, App
21	653	100.0	915	12	US-10-144-678A-834 Sequence 834, App
22	653	100.0	945	13	US-09-012-896-834 Sequence 1861, App
23	653	100.0	945	10	US-09-736-457-1861 Sequence 1861, App
24	653	100.0	945	10	US-09-902-941-1861 Sequence 1861, App
25	653	100.0	945	10	US-09-849-626-1861 Sequence 1861, App
26	653	100.0	945	14	US-10-017-754-1861 Sequence 1861, App
27	653	100.0	1012	9	US-09-735-705-351 Sequence 351, App
28	653	100.0	1012	10	US-09-850-716A-351 Sequence 351, App
29	653	100.0	1012	10	US-09-897-778-351 Sequence 351, App
30	653	100.0	1012	12	US-10-117-982-351 Sequence 351, App
31	653	100.0	1035	9	US-09-922-217-1084 Sequence 1084, App
32	653	100.0	1035	10	US-09-833-263-1084 Sequence 1084, App
33	653	100.0	1035	11	US-09-938-864-388 Sequence 388, App
34	653	100.0	1035	13	US-10-025-380-1084 Sequence 388, App
35	653	100.0	1035	14	US-10-125-635A-388 Sequence 388, App
36	653	100.0	1035	14	US-10-002-603-388 Sequence 388, App
37	653	100.0	1155	10	US-09-902-941-1875 Sequence 1875, App
38	653	100.0	1155	10	US-09-849-626-1875 Sequence 1875, App
39	653	100.0	1155	14	US-10-017-754-1875 Sequence 851, App
40	653	100.0	1203	9	US-09-759-143-851 Sequence 851, App
41	653	100.0	1203	9	US-09-780-669-851 Sequence 851, App
42	653	100.0	1203	9	US-09-822-827-851 Sequence 851, App
43	653	100.0	1203	10	US-09-895-793-851 Sequence 851, App
44	653	100.0	1203	10	US-09-895-814-851 Sequence 851, App
45	653	100.0	1203	12	US-10-144-678A-851 Sequence 851, App

#### ALIGNMENTS

#### RESULT 1

US-09-759-143-822

Sequence 822, Application US/09759143

Patent No. US200202248A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C23

; CURRENT APPLICATION NUMBER: US/09/759,143  
; CURRENT FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-759-143-822

Alignment Scores:  
Pred. No.: 7,59e-71 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-759-143-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCCTCGGCTTGGGTGTGTCGACAAACGACGCGGCGACGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTATCACCGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCACTCGGTGACCTGGCAACCAAGTCCGGGGCGACCGGTACAGGGAACGTG 381
DB 382 ACATTGGCGGAGGACCGCGGCC 405
```

## RESULT 2

US-09-780-669-822  
; Sequence 822, Application US/09780669  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT APPLICATION NUMBER: US/09/780,669  
; CURRENT FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-822

Alignment Scores:  
Pred. No.: 7,59e-71 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-780-669-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCCTCGGCTTGGGTGTGTCGACAAACGACGCGGCGACGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTATCACCGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCACTCGGTGACCTGGCAACCAAGTCCGGGGCGACCGGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGGAGGACCGCGGCC 405
```

## RESULT 3

US-09-822-827-822  
; Sequence 822, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.534C1  
; CURRENT APPLICATION NUMBER: US/09/822,827  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-827-822

Alignment Scores:  
Pred. No.: 7,59e-71 Length: 675  
Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9 Indels: 0  
DB: Gaps: 0

US-09-684-215A-18 (1-128) x US-09-822-827-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGCGCGCTCCGATACCTCCAGCTGCCAGGTGGCAGGATTCGCATTCGCATC 81
QY 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGCAGCGGATGGCGATCGCGGCCAGATCAAGCTCCACCGTTCATATCGGCGCTACC 141
QY 41 AlapheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGCTGGGTGTGTGCGACAACAGCGGCGGCGGCGGCGGCGGCGG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGATCAACTCGCGCCAGCGATCTCCACCGGCGGCGGCGGCGG 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGGCGCTCCGATCAACTCGCGCCAGCGGATCGCGGCGGCGGCGGCGGCGGCGG 321
QY 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGCGGCGGCGG 381
```

## RESULT 4

US-09-895-793-822

; Sequence 822, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-822

Alignment Scores:

Pred. No.: 7,59e-71 Length: 675

Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-895-793-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGCGCGCTCCGATACCTCCAGCTGCCAGGTGGCAGGATTCGCATTCGCATC 81
QY 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGCAGCGGATGGCGATCGCGGCCAGATCAAGCTCCACCGTTCATATCGGCGCTACC 141
QY 41 AlapheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGCTGGGTGTGTGCGACAACAGCGGCGGCGGCGGCGGCGGCGG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGGCGGCGGCGGCGG 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGGCGCTCCGATCAACTCGCGCCAGCGGATCGCGGCGGCGGCGGCGGCGGCGG 321
QY 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGCGGCGGCGG 381
```

## RESULT 5

US-09-895-814-822

; Sequence 822, Application US/09895814

; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895,814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-822

Alignment Scores:
Pred. No.: 7 59e-71 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-895-814-822 (1-675)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
DB 22 ACGGCGCGGTCGATAAATTCAGCTGTCCAGGTGGCAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATCGCATCCGCGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGTCGTGGTGTTCGACAAACACGACGCGGCGCAGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCCGAGCGCTCCCGCGCAAGTCTCGGCATCTCCACGCGGAGTGATCACCGGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGCCACCGGATGGCGGCGCGCTTAACGGGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 6
US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121 427C28
; CURRENT APPLICATION NUMBER: US/10/144, 678A
; NUMBER OF SEQ ID NOS: 11033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-678A-822

Alignment Scores:
Pred. No.: 7 59e-71 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-18 (1-128) x US-10-144-678A-822 (1-675)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
DB 22 ACGGCGCGGTCGATAAATTCAGCTGTCCAGGTGGCAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATCGCATCCGCGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGTCGTGGTGTTCGACAAACACGACGCGGCGCAGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCCGAGCGCTCCCGCGCAAGTCTCGGCATCTCCACGCGGAGTGATCACCGGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGCCACCGGATGGCGGCGCGCTTAACGGGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 7
US-10-012-896-822
; Sequence 822, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
```



```
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-822

Alignment Scores:
Pred. No.: 7,59e-71 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13

US-09-684-215A-18 (1-128) x US-10-012-896-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyClnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlnGlyGlnGlyPheAlaIleProIle 60
Db 142 GCCTTCCTCGCTGGGTGGTGTGACAAACACGCGCAGCGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 CTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGTGTATCACC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGACGTCG 381

RESULT 8
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriack
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
```

```
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:
Pred. No.: 9,73e-71 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10

US-09-684-215A-18 (1-128) x US-09-736-457-1862 (1-822)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyClnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlnGlyGlnGlyPheAlaIleProIle 60
Db 142 GCCTTCCTCGCTGGGTGGTGTGACAAACACGCGCAGCGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 CTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGTGTATCACC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGACGTCG 381

RESULT 9
US-09-902-941-1862
; Sequence 1862, Application US/09902941
; Patent No. US2002017952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darriack
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902.941
; CURRENT FILING DATE: 2001-07-10
```

; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1862  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-1862

Alignment Scores:  
Pred. No.: 9,73e-71 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-902-941-1862 (1-822)

```
QY 1 ThrAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
    |||||
DB 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
    |||||
DB 82 GGCAGGCGATGGCGATCCGCGCCAGATCAAGCTTCCACCCTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
    |||||
DB 142 GCCTTCTCGGCTTGGGTGTTCTCGACAAACGACGCGGCGAGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
    |||||
DB 202 GTCGGAGCGCTCCGCGCAAGTCTCGCATCTCCACGGCGAGTGTATCACC GGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
    |||||
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGGAGCGCTTAAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120
    |||||
DB 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
    |||||
DB 382 ACATTGGCGAGGACCCCGGCC 405
```

## RESULT 10

US-09-849-626-1862  
; Sequence 1862, Application US/09849626  
; Publication No. US2002019769A1  
; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya  
; APPLICANT: Fanger, Gary  
; APPLICANT: Wang, Aljun  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Switzer, Anne  
; APPLICANT: McNeill, Patricia  
; APPLICANT: Clapper, Jonathan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C16  
; CURRENT APPLICATION NUMBER: US/09/849,626  
; CURRENT FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 1926  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1862  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-849-626-1862

Alignment Scores:  
Pred. No.: 9,73e-71 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
DB: 10 Gaps: 0

Score: 653.00 Matches: 128  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-849-626-1862 (1-822)

```
QY 1 ThrAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
    |||||
DB 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
    |||||
DB 82 GGCAGGCGATGGCGATCCGCGCCAGATCAAGCTTCCACCCTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
    |||||
DB 142 GCCTTCTCGGCTTGGGTGTTCTCGACAAACGACGCGGCGAGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
    |||||
DB 202 GTCGGAGCGCTCCGCGCAAGTCTCGCATCTCCACGGCGAGTGTATCACC GGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
    |||||
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGGAGCGCTTAAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120
    |||||
DB 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
    |||||
DB 382 ACATTGGCGAGGACCCCGGCC 405
```

## RESULT 11

US-10-017-754-1862  
; Sequence 1862, Application US/10017754  
; Publication No. US20030054363A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C18  
; CURRENT APPLICATION NUMBER: US/10/017,754  
; CURRENT FILING DATE: 2001-10-29  
; NUMBER OF SEQ ID NOS: 2004  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1862  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-754-1862

Alignment Scores:  
Pred. No.: 9,73e-71 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-18 (1-128) x US-10-017-754-1862 (1-822)

```
Oy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCATTCGCATC 81
Oy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
Oy 41 AlaphelLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGCTTGGGTGTGTCGACAAACACGCGCAGCGCAGTCCACGCGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 142 GCCTTCCTCGCTTGGGTGTGTCGACAAACACGCGCAGCGCAGTCCACGCGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGTGCATCCCGCGTC 261
Oy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAACGGGCGATCATCC 321
Oy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAACGTCGGCGCGCAGCGGTACAGGGAACGTG 381
Oy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405
```

#### RESULT 12

US-09-735-705-353

; Sequence 353, Application US/09735705

; Patent No. US20020052329A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Lihun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C14

; CURRENT APPLICATION NUMBER: US/09/735,705

; CURRENT FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 419

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 353

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-735-705-353

Alignment Scores:

Pred. No.: 1.09e-70 Length: 900

Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-735-705-353 (1-900)

```
Oy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
```

```
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCATTCGCATC 81
```

```
Oy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
Oy 41 AlaphelLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGCTTGGGTGTGTCGACAAACACGCGCAGCGCAGTCCACGCGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGTGCATCCCGCGTC 261
Oy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAACGGGCGATCATCC 321
Oy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAACGTCGGCGCGCAGCGGTACAGGGAACGTG 381
Oy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405
```

#### RESULT 13

US-09-850-716A-353

; Sequence 353, Application US/09850716A

; Patent No. US20020115139A1

; GENERAL INFORMATION:

; APPLICANT: Kalos, Michael D.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Retter, Marc W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C15

; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 353

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-850-716A-353

Alignment Scores:

Pred. No.: 1.09e-70 Length: 900

Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-850-716A-353 (1-900)

```
Oy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
```

```
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTCGCATTCGCATC 81
```

```
Oy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
```

```
Db 82 GGCAGCGGATGGCGATCGGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
```

```
Oy 41 AlaphelLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
```

```
Db 142 GCCTTCCTCGCTTGGGTGTGTCGACAAACACGCGCAGCGCAGTCCACGCGTG 201
```

```
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
```

```
Db 202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGTGCATCCCGCGTC 261
```

```
Oy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
```

Db 262 GACGGCGCTCCGATCAACTCGGCCACCGGATCGCGGACGCGCTTAACGGGATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGGTGACCTGGCNAACCAAGTGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProla 128  
Db 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 14

US-09-897-778-353  
; Sequence 353, Application US/0989778  
; Patent No. US20020147143A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Warnerakis, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Peckham, David W.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C16  
; CURRENT APPLICATION NUMBER: US/09/897,778  
; NUMBER OF SEQ ID NOS: 467  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens

## US-09-897-778-353

Alignment Scores:  
Pred. No.: 1.09e-70 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-897-778-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATCGCATCGCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGTGGTGTTCGACAAACACGCAACGCGGCGAGTCCCAACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTGCATACCGCGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGCGGATGGCGGAGCGCTTAACGGGATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGGTGACCTGGCNAACCAAGTGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProla 128

Search completed: September 5, 2003, 12:04:47  
Job time: 189.154 secs

Db 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 15

US-10-117-982-353  
; Sequence 353, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210120.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens

## US-10-117-982-353

Alignment Scores:  
Pred. No.: 1.09e-70 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-684-215A-18 (1-128) x US-10-117-982-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATCGCATCGCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGTGGTGTTCGACAAACACGCAACGCGGCGAGTCCCAACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTGCATACCGCGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGCGGATGGCGGAGCGCTTAACGGGATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGGTGACCTGGCNAACCAAGTGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProla 128  
Db 382 ACATTGGCGGAGGACCCCGGCC 405

Search completed: September 5, 2003, 12:04:47  
Job time: 189.154 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:54:40 ; Search time 43.5517 Seconds  
(without alignments)  
1584.308 Million cell updates/sec

Title: US-09-684-215A-17

Perfect score: 148

Sequence: 1 TAAADNFQLSQGGGFAIPGQAMAIAGQI 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool.p/US09684215/runat\_05092003\_072203\_619/app.query.fasta\_1.853  
-DB=Published Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09684215 -CGN\_1\_130 @runat\_05092003\_072203\_619  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	ID	Description
1	148	100.0	186	10 US-09-902-941-1879

2	148	100.0	186	10 US-09-849-626-1879	Sequence 1879, Ap
3	148	100.0	186	14 US-10-017-754-1879	Sequence 1879, Ap
4	148	100.0	399	9 US-09-757-417-46	Sequence 46, Appl
5	148	100.0	399	14 US-10-042-945-46	Sequence 46, Appl
6	148	100.0	447	12 US-10-084-843-4	Sequence 4, Appl
7	148	100.0	447	12 US-10-193-002-4	Sequence 4, Appl
8	148	100.0	675	9 US-09-759-143-822	Sequence 822, App
9	148	100.0	675	9 US-09-780-669-822	Sequence 822, App
10	148	100.0	675	9 US-09-822-827-822	Sequence 822, App
11	148	100.0	675	10 US-09-895-793-822	Sequence 822, App
12	148	100.0	675	10 US-09-895-814-822	Sequence 822, App
13	148	100.0	675	12 US-10-144-678A-822	Sequence 822, App
14	148	100.0	675	13 US-10-012-896-822	Sequence 822, App
15	148	100.0	702	9 US-09-287-849-27	Sequence 27, Appl
16	148	100.0	702	12 US-10-359-460-27	Sequence 27, Appl
17	148	100.0	822	10 US-09-736-457-1862	Sequence 1862, Ap
18	148	100.0	822	10 US-09-902-941-1862	Sequence 1862, Ap
19	148	100.0	822	10 US-09-849-626-1862	Sequence 1862, Ap
20	148	100.0	822	14 US-10-017-754-1862	Sequence 1862, Ap
21	148	100.0	861	10 US-09-902-941-1877	Sequence 1877, Ap
22	148	100.0	861	10 US-09-849-626-1877	Sequence 1877, Ap
23	148	100.0	861	14 US-10-017-754-1877	Sequence 1877, Ap
24	148	100.0	900	9 US-09-735-705-353	Sequence 353, App
25	148	100.0	900	10 US-09-850-716A-353	Sequence 353, App
26	148	100.0	900	10 US-09-897-778-353	Sequence 353, App
27	148	100.0	900	12 US-10-117-982-353	Sequence 353, App
28	148	100.0	915	9 US-09-759-143-834	Sequence 834, App
29	148	100.0	915	9 US-09-780-669-834	Sequence 834, App
30	148	100.0	915	9 US-09-822-827-834	Sequence 834, App
31	148	100.0	915	10 US-09-895-793-834	Sequence 834, App
32	148	100.0	915	10 US-09-895-814-834	Sequence 834, App
33	148	100.0	915	12 US-10-144-678A-834	Sequence 834, App
34	148	100.0	915	13 US-10-012-896-834	Sequence 834, App
35	148	100.0	945	10 US-09-736-457-1861	Sequence 1861, Ap
36	148	100.0	945	10 US-09-902-941-1861	Sequence 1861, Ap
37	148	100.0	945	10 US-09-849-626-1861	Sequence 1861, Ap
38	148	100.0	945	14 US-10-017-754-1861	Sequence 1861, Ap
39	148	100.0	1012	9 US-09-735-705-351	Sequence 351, App
40	148	100.0	1012	10 US-09-850-716A-351	Sequence 351, App
41	148	100.0	1012	10 US-09-897-778-351	Sequence 351, App
42	148	100.0	1012	12 US-10-117-982-351	Sequence 351, App
43	148	100.0	1035	9 US-09-922-217-1084	Sequence 1084, Ap
44	148	100.0	1035	10 US-09-833-263-1084	Sequence 1084, Ap
45	148	100.0	1035	11 US-09-938-864-388	Sequence 388, App

#### ALIGNMENTS

RESULT 1  
US-09-902-941-1879  
; Sequence 1879, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnierakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vadnick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902.941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1879  
; LENGTH: 186

10



```
Db 22 ACGCGCCGTCGCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCGCATC 81
Qy 21 GlycAlaMetAlaIleAlaGlyGlnIle 30
    |||||
Db 82 GGCAGCGGATGGCGATCGCGGCCAGATC 111

RESULT 5
US-10-042-945-46
; Sequence 46, Application US/10042945
; Publication No. US20030045468A1
; GENERAL INFORMATION:
; APPLICANT: Fling, Steven P.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Wang, AiJun
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY,
; FILE OF INVENTION: DIAGNOSIS AND MONITORING OF BREAST CANCER
; FILE REFERENCE: 210121.479C3
; CURRENT APPLICATION NUMBER: US/10/042,945
; CURRENT FILING DATE: 2002-01-08
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-945-46

Alignment Scores:
Pred. No.: 2,44e-16 Length: 399
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x US-10-042-945-46 (1-399)
Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCGCCGTCGCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCGCATC 81
Qy 21 GlycAlaMetAlaIleAlaGlyGlnIle 30
    |||||
Db 82 GGCAGCGGATGGCGATCGCGGCCAGATC 111

RESULT 6
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4

Alignment Scores:
Pred. No.: 2,78e-16 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-17 (1-30) x US-10-084-843-4 (1-447)
Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
    |||||
Db 11 ACGCGCCGTCGCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCGCATC 70
Qy 21 GlycAlaMetAlaIleAlaGlyGlnIle 30
    |||||
Db 71 GGCAGCGGATGGCGATCGCGGCCAGATC 100

RESULT 7
US-10-193-002-4
; Sequence 4, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/193,002  
FILING DATE: 10-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-193-002-4

Alignment Scores:  
Pred. No.: 2,78e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps:

US-09-684-215A-17 (1-30) x US-10-193-002-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db II ACGGCGCGGTCCGATAACTTCCAGCTGTCAGGTTGGCAGGGATTCCGCATTCGGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATC 100

## RESULT 8

US-09-759-143-822  
; Sequence 822, Application US/09759143  
; Patent No. US200202248A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT FILING DATE: 2001-01-12  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps:

US-09-684-215A-17 (1-30) x US-09-780-669-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||

; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-759-143-822  
Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps:

US-09-684-215A-17 (1-30) x US-09-759-143-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 22 ACGGCGCGGTCCGATAACTTCCAGCTGTCAGGTTGGCAGGGATTCCGCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 82 GGGCAGCGCATGGCGATCGCGGCCAGATC 111

## RESULT 9

US-09-780-669-822  
; Sequence 822, Application US/09780669  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT FILING DATE: 2001-02-09  
; CURRENT APPLICATION NUMBER: US/09/780,669  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps:

Db 22 ACGGCCGCGTCCGATACTTCCAGCTGTCCAGGGTGGGAGGGATTGCCATTCCGATC 81

Qy 21 GlyAlaMetAlaIleAlaGlyGlnIle 30  
|||||

Db 82 GGGCAGCGGATGGGATCGCGGGCCAGATC 111

## RESULT 10

US-09-822-827-822

; Sequence 822, Application US/09822827

; Patent No. US20020081680A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C1

; CURRENT APPLICATION NUMBER: US/09/822,827

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-822-827-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-822-827-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||

Db 22 ACGGCCGCGTCCGATACTTCCAGCTGTCCAGGGTGGGAGGGATTGCCATTCCGATC 81

Qy 21 GlyAlaMetAlaIleAlaGlyGlnIle 30  
|||||

Db 82 GGGCAGCGGATGGGATCGCGGGCCAGATC 111

## RESULT 11

US-09-895-793-822

; Sequence 822, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-895-793-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||

Db 22 ACGGCCGCGTCCGATACTTCCAGCTGTCCAGGGTGGGAGGGATTGCCATTCCGATC 81

Qy 21 GlyAlaMetAlaIleAlaGlyGlnIle 30  
|||||

Db 82 GGGCAGCGGATGGGATCGCGGGCCAGATC 111

## RESULT 12

US-09-895-814-822

; Sequence 822, Application US/09895814

; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895,814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-814-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-895-814-822 (1-675)

QY 1 ThrAlaAlaSerAspAspPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
|||||  
DB 22 ACGGCGGCTCCGATAACTTCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGCATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
DB 82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

## RESULT 13

US-10-144-678A-822

; Sequence 822, Application US/10144678A

; Publication No. US2003015708A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Derrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals y de Bassols, Carlota  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Deng, Ta

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C28

; CURRENT APPLICATION NUMBER: US/10144, 678A

; CURRENT FILING DATE: 2002-08-12

; NUMBER OF SEQ ID NOS: 1033

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-144-678A-822

Alignment Scores:

Pred. No.: 4,47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-684-215A-17 (1-30) x US-10-144-678A-822 (1-675)

QY 1 ThrAlaAlaSerAspAspPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
|||||  
DB 22 ACGGCGGCTCCGATAACTTCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGCATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
DB 82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

## RESULT 14

US-10-012-896-822

; Sequence 822, Application US/10012896

; Publication No. US2002083251A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Derrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Meagher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C27  
; CURRENT APPLICATION NUMBER: US/10/012, 896  
; CURRENT FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 1011  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-012-896-822

Alignment Scores:

Pred. No.: 4,47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-684-215A-17 (1-30) x US-10-012-896-822 (1-675)

QY 1 ThrAlaAlaSerAspAspPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
|||||  
DB 22 ACGGCGGCTCCGATAACTTCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGCATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
DB 82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

## RESULT 15

US-09-287-849-27

; Sequence 27, Application US/09287849

; Patent No. US20020009459A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Nejo, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; AND THEIR USES

```

; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

```

```

Alignment Scores:
Pred. No.: 4.68e-16 Length: 702
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-09-684-215A-17 (1-30) x US-09-287-849-27 (1-702)

```

Qy 1 ThrAlaAlaSerAspPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 25 ACGGCGGTCGCGATAACTTCCAGCTGCCAGGTTGGCAGGGATTCCGCATTCGCATC 84
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 85 GGGCAGCGGATGGCGATCGCGGGCCAGATC 114

```

Search completed: September 5, 2003, 12:04:44  
Job time : 44.8851 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:52:39 ; Search time 535.655 seconds  
(without alignments)  
2291.192 Million cell updates/sec

Title: US-09-684-215A-17  
Perfect score: 148  
Sequence: 1 TAASDNFQLSQGGQGAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2.1/USPTO.spool\_p/US09684215/runat\_05092003\_072202\_559/app\_query.fasta\_1.853  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=Dits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcet -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 @cgn.1.1.3962 @runat\_05092003\_072202\_559 -NCPU=3  
-NO\_MMAB -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	148	100.0	186	6	AX369169 Sequence
2	148	100.0	447	6	ARI69152 Sequence
3	148	100.0	447	6	ARI82442 Sequence
4	148	100.0	447	6	ARI194825 Sequence
5	148	100.0	447	6	AR233097 Sequence
6	148	100.0	447	6	AX429596 Sequence
7	148	100.0	447	6	BD006325 Compounds
8	148	100.0	447	6	BD006445 Compounds
9	148	100.0	447	6	BD069285 Compounds
10	148	100.0	675	6	AR261272 Sequence
11	148	100.0	675	6	AX201049 Sequence
12	148	100.0	675	6	AX267848 Sequence
13	148	100.0	822	6	AR277645 Sequence
14	148	100.0	822	6	AX369152 Sequence
15	148	100.0	861	6	AX369167 Sequence
16	148	100.0	894	6	AX351489 Sequence
17	148	100.0	900	6	AR220690 Sequence
18	148	100.0	900	6	AX365960 Sequence
19	148	100.0	915	6	AR261281 Sequence
20	148	100.0	915	6	AX201061 Sequence
21	148	100.0	915	6	AX267860 Sequence
22	148	100.0	945	6	AR277644 Sequence
23	148	100.0	945	6	AX369151 Sequence
24	148	100.0	1012	6	AR220689 Sequence
25	148	100.0	1012	6	AX365958 Sequence
26	148	100.0	1068	6	AX005788 Sequence
27	148	100.0	1143	6	AX005790 Sequence
28	148	100.0	1155	6	AX369165 Sequence
29	148	100.0	1203	6	AX201078 Sequence
30	148	100.0	1203	6	AX267877 Sequence
31	148	100.0	1464	6	AR229410 Sequence
32	148	100.0	1464	6	AX156105 Sequence
33	148	100.0	1464	6	AX361955 Sequence
34	148	100.0	1557	6	AR229398 Sequence
35	148	100.0	1557	6	AX156089 Sequence
36	148	100.0	1557	6	AX361939 Sequence
37	148	100.0	1578	6	AR229328 Sequence
38	148	100.0	1578	6	AX155945 Sequence
39	148	100.0	1578	6	AX361795 Sequence
40	148	100.0	1590	6	AX316986 Sequence
41	148	100.0	1752	6	AR229413 Sequence
42	148	100.0	1752	6	AX156109 Sequence
43	148	100.0	1752	6	AX361959 Sequence
44	148	100.0	1758	6	AR229401 Sequence
45	148	100.0	1758	6	AX156093 Sequence

ALIGNMENTS

```
AX369169
LOCUS AX369169 186 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1879 from Patent WO0204514.
ACCESSION AX369169
VERSION AX369169.1 GI:18857178
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
Mannerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 1879 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..186
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 39 a 61 c 47 g 39 t
ORIGIN
Alignment Scores:
Pred. No.: 3,84e-14 Length: 186
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX369169 (1-186)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATC 111

RESULT 2
AR169152
LOCUS AR169152 447 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 4 from patent US 6290969.
ACCESSION AR169152
VERSION AR169152.1 GI:17906927
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,
Vedvick, T.S. and Twardzik, D.R.
TITLE Compositions and methods for immunotherapy and diagnosis of
tuberculosis
JOURNAL Patent: US 6290969-A 4 18-SEP-2001;
FEATURES
source
Location/Qualifiers
1..447
/organism="unknown"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 9,83e-14 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR169152 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100

RESULT 4
AR194825
LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6350456.
ACCESSION AR194825
VERSION AR194825.1 GI:20244262
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed, S.G., Skeiky, Y.A.W. and Dillon, D.C.
TITLE Compositions and methods for the prevention and treatment of M.
tuberculosis infection
JOURNAL Patent: US 6350456-A 4 26-FEB-2002;
FEATURES
source
Location/Qualifiers
1..447
/organism="unknown"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 9,83e-14 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR194825 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100

RESULT 3
AR182442
LOCUS AR182442 447 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6338852.
ACCESSION AR182442
VERSION AR182442.1 GI:20225649
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,
Vedvick, T.S. and Twardzik, D.R.
TITLE Compounds and methods for diagnosis of tuberculosis
JOURNAL Patent: US 6338852-A 4 15-JAN-2002;
FEATURES
source
Location/Qualifiers
1..447
/organism="unknown"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 9,83e-14 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR182442 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100
```



Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR194825 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 11 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGCATC 70  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGCGGATGGCGATCGCGGGCCAGATC 100

## RESULT 5

AR233097 AR233097 447 bp DNA linear PAT 20-DEC-2002  
LOCUS  
DEFINITION Sequence 4 from patent US 6458366.  
ACCESSION AR233097  
VERSION AR233097.1 GI:27275533  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,  
Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.

TITLE Compounds and methods for diagnosis of tuberculosis  
JOURNAL Patent: US 6458366-A 4 01-OCT-2002;  
FEATURES Location/Qualifiers  
source 1. .447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR233097 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 11 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGCATC 70  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGCGGATGGCGATCGCGGGCCAGATC 100

## RESULT 6

AX429596 AX429596 447 bp DNA linear PAT 21-JUN-2002  
LOCUS  
DEFINITION Sequence 4 from Patent EP1203817.  
ACCESSION AX429596  
VERSION AX429596.1 GI:21540845  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE 1  
AUTHORS Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,  
Vedvick,T.S. and Twardzik,D.R.

TITLE Compounds and methods for immunotherapy and diagnosis of  
JOURNAL tuberculosis  
Patent: EP 1203817-A 4 08-MAY-2002;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers

source 1. .447  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX429596 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 11 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGCATC 70  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGCGGATGGCGATCGCGGGCCAGATC 100

## RESULT 7

BD006325 BD006325 447 bp DNA linear PAT 31-JAN-2002  
LOCUS  
DEFINITION Compounds and methods for diagnosis of Tuberculosis.

ACCESSION BD006325  
VERSION BD006325.1 GI:18634696  
KEYWORDS JP 2001500383-A/4.  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,  
Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.  
TITLE Compounds and methods for diagnosis of Tuberculosis  
JOURNAL Patent: JP 2001500383-A 4 16-JAN-2001;  
CORIXA CORP

COMMENT OS Unidentified  
PN JP 2001500383-A/4  
PD 16-JAN-2001  
PF 07-OCT-1997 JP 1998518432  
PR 11-OCT-1996 US 08/729622,13-MAR-1997 US 08/818111 PI

STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS  
NETO,  
PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI  
MICHAEL J LODES  
PC C12N15/31.C07K14/35.C07K16/12.C12Q1/68.C12N15/62.G01N33/53 CC  
Strandedness: Single;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT source 1. .447  
FT /organism="Unidentified".

FEATURES  
source 1. .447  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD006325 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGCATC 70  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 71 GGGCAGCGATGGCGATCGCGGCCAGATC 100

RESULT 8  
 BD006445  
 LOCUS 447 bp DNA linear PAT 31-JAN-2002  
 DEFINITION Compounds and methods for immunotherapy and diagnosis of  
 Tuberculosis.  
 ACCESSION BD006445  
 VERSION BD006445.1 GI:18634816  
 KEYWORDS JP 2001501832-A/4.  
 SOURCE unidentified  
 ORGANISM unclassified  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,  
 Vedvick, T.S., Twardzik, D.R. and Lodes, M.J.  
 TITLE Compounds and methods for immunotherapy and diagnosis of  
 JOURNAL Tuberculosis.  
 COMMENT Patent: JP 2001501832-A 4 13-FEB-2001;  
 CORIXA CORP  
 OS Unidentified  
 PN JP 2001501832-A/4  
 PD 13-FEB-2001  
 PF 07-OCT-1997 JP 1998518456  
 PR 11-OCT-1996 US 08/730510, 13-MAR-1997 US 08/818112 PI  
 STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS  
 NETO,  
 PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI  
 MICHAEL J LODES  
 PC C12N15/31, C07K14/35, A61K39/04, A61K48/00, A61K49/00, C12N15/62,  
 PC C07K19/00,  
 PC G01N33/50, G01N33/60, G01N33/569, C12N1/19, C12N1/20, C12N1/21, PC  
 C12N5/10//  
 PC (C12N1/21, C12R1:19)  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers  
 FT source  
 FT 1. .447  
 Location/Qualifiers  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 79 a 146 c 149 g 72 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9,83e-14 Length: 447  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD006445 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGCATC 70  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 71 GGGCAGCGATGGCGATCGCGGCCAGATC 100

RESULT 9  
 BD069285  
 LOCUS 675 bp DNA linear PAT 29-JAN-2003  
 DEFINITION Sequence 822 from patent US 6321716.  
 ACCESSION AR261272  
 VERSION AR261272.1 GI:28072035  
 KEYWORDS Unknown.  
 SOURCE Unclassified.  
 REFERENCE 1 (bases 1 to 675)  
 AUTHORS Mashiki, Z., and Harada, J.

LOCUS 447 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Compounds and methods for immunotherapy and diagnosis of  
 tuberculosis.  
 ACCESSION BD069285  
 VERSION BD069285.1 GI:22614888  
 KEYWORDS JP 2001517069-A/4.  
 SOURCE unidentified  
 ORGANISM unclassified  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,  
 Vedvick, T.H. and Twardzik, D.R.  
 TITLE Compounds and methods for immunotherapy and diagnosis of  
 JOURNAL Tuberculosis.  
 COMMENT Patent: JP 2001517069-A 4 02-OCT-2001;  
 CORIXA CORP  
 OS Unidentified  
 PN JP 2001517069-A/4  
 PD 02-OCT-2001  
 PF 30-AUG-1996 JP 1997511464  
 PR 01-SEP-1995 US 08/523436, 22-SEP-1995 US 08/533634 PR  
 22-MAR-1996 US 08/620874, 05-JUN-1996 US 08/659683 PR  
 12-JUL-1996 US 08/680574  
 PI STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, ANTONIO CAMPOS  
 PI NETO,  
 PI RAYMOND HOUGHTON, THOMAS H VEDVICK, DANIEL R TWARDZIK PC  
 C12N15/31, C07K14/35, A61K38/16, C12N15/62, G01N33/569, C12Q1/68, PC  
 C12N5/10,  
 PC C12N1/21//A61K39/04, C12N1/21, C12R1:19)  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 CC Compounds and methods for immunotherapy and diagnosis of CC  
 tuberculosis  
 FH Key Location/Qualifiers  
 FT source  
 FT 1. .447  
 Location/Qualifiers  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 79 a 146 c 149 g 72 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9,83e-14 Length: 447  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD069285 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGCATC 70  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 71 GGGCAGCGATGGCGATCGCGGCCAGATC 100

RESULT 10  
 AR261272  
 LOCUS 675 bp DNA linear PAT 29-JAN-2003  
 DEFINITION Sequence 822 from patent US 6321716.  
 ACCESSION AR261272  
 VERSION AR261272.1 GI:28072035  
 KEYWORDS Unknown.  
 SOURCE Unclassified.  
 REFERENCE 1 (bases 1 to 675)  
 AUTHORS Mashiki, Z., and Harada, J.

TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 822 27-NOV-2001;

FEATURES  
source  
1. .675  
Location/Qualifiers

BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.53e-13 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR261272 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCCGATC 81  
|||||  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 82 GGCAGCGGATGCGATCGCGGCCAGATC 111  
|||||

RESULT 11

AX201049  
LOCUS AX201049 675 bp DNA linear PAT 29-AUG-2001

DEFINITION Sequence 679 from Patent WO0151633.

ACCESSION AX201049

VERSION AX201049.1 GI:15390857

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W.,  
Stolk, J.A., Skeiky, Y.A., Wang, A. and Weagher, M.J.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0151633-A 679 19-JUL-2001;

LOCUS AX201049 675 bp DNA linear PAT 29-AUG-2001

FEATURES  
source  
1. .675  
Location/Qualifiers

BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.53e-13 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX201049 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||

Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCCGATC 81  
|||||

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||

Db 82 GGCAGCGGATGCGATCGCGGCCAGATC 111  
|||||

RESULT 12

AX267848

LOCUS AX267848 822 bp DNA linear PAT 26-OCT-2001

DEFINITION Sequence 822 from Patent WO0173032.

ACCESSION AX267848

VERSION AX267848.1 GI:16516494

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,

Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,

Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T.

and Henderson, R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0173032-A 822 04-OCT-2001;

LOCUS AX267848 822 bp DNA linear PAT 26-OCT-2001

DEFINITION Sequence 822 from Patent WO0173032.

ACCESSION AX267848

VERSION AX267848.1 GI:16516494

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,

Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,

Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T.

and Henderson, R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0173032-A 822 04-OCT-2001;

LOCUS AX267848 822 bp DNA linear PAT 26-OCT-2001

DEFINITION Sequence 822 from Patent WO0173032.

ACCESSION AX267848

VERSION AX267848.1 GI:16516494

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,

Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,

Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T.

and Henderson, R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0173032-A 822 04-OCT-2001;

LOCUS AX267848 822 bp DNA linear PAT 26-OCT-2001

DEFINITION Sequence 822 from Patent WO0173032.

ACCESSION AX267848

VERSION AX267848.1 GI:16516494

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,

Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,

Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T.

and Henderson, R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0173032-A 822 04-OCT-2001;

US-09-684-215A-17 (1-30) x AR277645 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 22 ACGGCGCGGTCCGATAACTTCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

RESULT 14  
 AX369152  
 LOCUS 822 bp DNA linear PAT 16-FEB-2002  
 DEFINITION Sequence 1862 from Patent WO204514.  
 ACCESSION AX369152  
 VERSION AX369152.1 GI:18857170  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1  
 REFERENCE  
 AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,  
 Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,  
 McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.  
 TITLE Compositions and methods for the therapy and diagnosis of lung  
 cancer  
 JOURNAL Patent: WO 0204514-A 1862 17-JAN-2002;  
 CORIXA CORPORATION (US)  
 FEATURES  
 source 1. 822  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 146 a 281 c 270 g 125 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.89e-13 Length: 822  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX369152 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 22 ACGGCGCGGTCCGATAACTTCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

RESULT 15  
 AX369167  
 LOCUS 861 bp DNA linear PAT 16-FEB-2002  
 DEFINITION Sequence 1877 from Patent WO204514.  
 ACCESSION AX369167  
 VERSION AX369167.1 GI:18857177  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1  
 REFERENCE  
 AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,  
 Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,  
 McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.  
 TITLE Compositions and methods for the therapy and diagnosis of lung  
 cancer

JOURNAL Patent: WO 0204514-A 1877 17-JAN-2002;  
 CORIXA CORPORATION (US)  
 FEATURES  
 source 1. 861  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 186 a 244 c 209 g 222 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.98e-13 Length: 861  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX369167 (1-861)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 22 ACGGCGCGGTCCGATAACTTCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

Search completed: September 5, 2003, 10:31:40  
 Job time : 536.989 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:28:59 ; Search time 16.0345 Seconds  
(without alignments)  
825.814 Million cell updates/sec

Title: US-09-684-215A-17

Perfect score: 148

Sequence: 1 TAASDNFQLSQGGGFAIPICQAMAIAGQI 30

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_593/app\_query.fasta\_1.853  
-DB-Issued\_Patents\_NA -QPMF=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -CGN\_1\_145 @runat\_05092003\_072203\_593 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_1/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_1/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_1/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_1/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_1/ptodata/1/ina/PTCUS\_COMB.seq: \*  
6: /cgn2\_1/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	148	100.0	447	3	US-08-818-112-4
2	148	100.0	447	4	US-08-818-111-4
3	148	100.0	447	4	US-09-056-556-4
4	148	100.0	447	4	US-09-072-596-4
5	148	100.0	822	4	US-09-736-457-1862
6	148	100.0	900	4	US-09-643-597-353
7	148	100.0	900	4	US-09-606-421B-353
8	148	100.0	945	4	US-09-736-457-1861
9	148	100.0	1012	4	US-09-643-597-351
10	148	100.0	1012	4	US-09-606-421B-351
11	148	100.0	1464	4	US-09-620-412C-348
12	148	100.0	1464	4	US-09-598-419-348

Sequence 332, App  
Sequence 332, App  
Sequence 188, App  
Sequence 188, App  
Sequence 188, App  
Sequence 188, App  
Sequence 332, App  
Sequence 332, App  
Sequence 336, App  
Sequence 308, App  
Sequence 308, App  
Sequence 308, App  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 324, App  
Sequence 324, App  
Sequence 316, App  
Sequence 316, App  
Sequence 340, App  
Sequence 340, App  
Sequence 356, App  
Sequence 356, App  
Sequence 312, App  
Sequence 312, App  
Sequence 344, App  
Sequence 344, App  
Sequence 320, App  
Sequence 320, App  
Sequence 320, App  
Sequence 320, App  
Sequence 328, App  
Sequence 1, Appli  
Sequence 2, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-818-112-4  
Sequence 4, Application US/08818112  
Patent No. 6290959

#### GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 153  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,112  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C6  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-112-4

Alignment Scores:  
Pred. No.: 9.5e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-17 (1-30) x US-08-818-112-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATC 100

## RESULT 2

US-08-818-111-4

; Sequence 4, Application US/08818111

; Patent No. 6338852

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Campos-Nelo, Antonia

; APPLICANT: Houghton, Raymond

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Twardzik, Daniel R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS

; NUMBER OF SEQUENCES: 148

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/818,111

; FILING DATE: 13-Mar-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.417C6

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 447 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-818-111-4

## Alignment Scores:

Pred. No.: 9.5e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-08-818-111-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATC 100

## RESULT 3

US-09-056-556-4

; Sequence 4, Application US/09056556

; Patent No. 6350456

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

; NUMBER OF SEQUENCES: 241

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-Apr-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.457

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 447 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-09-056-556-4

## Alignment Scores:

Pred. No.: 9.5e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-056-556-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGCATC 70



Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-643-597-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
DB 22 ACGGGCGGTCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCCATTCGGATC 81  
|||||

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
DB 82 GGGCAGGCGATGCGGATCGCGGCCAGATC 111  
|||||

## RESULT 7

US-09-606-421B-353  
; Sequence 353, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Lijun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-606-421B-353

Alignment Scores:  
Pred. No.: 2.27e-15 Length: 900  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-606-421B-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
DB 22 ACGGGCGGTCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCCATTCGGATC 81  
|||||

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
DB 82 GGGCAGGCGATGCGGATCGCGGCCAGATC 111  
|||||

## RESULT 8

US-09-736-457-1861  
; Sequence 1861, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedwick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc

US-09-736-457-1861

Alignment Scores:  
Pred. No.: 2.63e-15 Length: 1012  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-643-597-351 (1-1012)

; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Lijun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1861  
; LENGTH: 945  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-736-457-1861

Alignment Scores:  
Pred. No.: 2.42e-15 Length: 945  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-736-457-1861 (1-945)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
DB 22 ACGGGCGGTCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCCATTCGGATC 81  
|||||

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
DB 82 GGGCAGGCGATGCGGATCGCGGCCAGATC 111  
|||||

## RESULT 9

US-09-643-597-351  
; Sequence 351, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Lijun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 351  
; LENGTH: 1012  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-643-597-351





```
Alignment Scores:
Pred. No.: 4.5e-15 Length: 1557
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-620-412C-332 (1-1557)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 82 GGCAGGCGATGCGGATCCGGGCCAGATC 111

RESULT 14
US-09-598-419-332
; Sequence 332, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 332
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-332

Alignment Scores:
Pred. No.: 4.5e-15 Length: 1557
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-598-419-332 (1-1557)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 82 GGCAGGCGATGCGGATCCGGGCCAGATC 111

RESULT 15
US-09-556-877-188
; Sequence 188, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305

; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 188
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-556-877-188

Alignment Scores:
Pred. No.: 4.57e-15 Length: 1578
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-556-877-188 (1-1578)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGGATTCCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 82 GGCAGGCGATGCGGATCCGGGCCAGATC 111

Search completed: September 5, 2003, 12:12:03
Job time : 20.0345 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:27:24 ; Search time 518.276 Seconds  
(without alignments)  
1406.846 Million cell updates/sec

Title: US-09-684-215A-17  
Perfect score: 148  
Sequence: 1 TAASDNFQLSQGGQGAIPIGQAMAIAGQI 30  
Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool.p/US09684215/runat\_05092003\_072202\_572/app\_query.fasta\_1.853  
-DB=EST\_QFMT-fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215.ecgn\_1.1.4575.erunat\_05092003\_072202\_572 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estinv:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
c 1	60	40.5	747	14	CB597827
c 2	57	38.5	238	10	BB077789
c 3	57	38.5	336	14	CB184974
4	57	38.5	343	14	CA930964
5	57	38.5	351	12	BI070561
6	57	38.5	366	12	BI055877
7	57	38.5	370	14	CA931030
c 8	57	38.5	370	14	CA931960
9	57	38.5	370	14	CA931997
10	57	38.5	370	14	CA932035
11	57	38.5	370	14	CA932108
12	57	38.5	370	14	CA932138
c 13	57	38.5	374	14	CA927770
c 14	57	38.5	374	14	CA928449
c 15	57	38.5	374	14	CA928493
c 16	57	38.5	374	14	CA928769
c 17	57	38.5	384	14	CA931116
c 18	57	38.5	385	14	CA931394
c 19	57	38.5	385	14	CA931501
20	57	38.5	387	14	CA931767
21	57	38.5	388	14	CA927636
c 22	57	38.5	388	14	CA928025
c 23	57	38.5	388	14	CA929028
c 24	57	38.5	389	14	CA927727
c 25	57	38.5	389	14	CA928142
26	57	38.5	390	14	CA927617
c 27	57	38.5	390	14	CA927744
c 28	57	38.5	390	14	CA927930
c 29	57	38.5	402	14	CA933902
c 30	57	38.5	403	13	BU824075
c 31	57	38.5	404	14	CA934515
c 32	57	38.5	408	14	CA930794
c 33	57	38.5	409	14	CA931302
c 34	57	38.5	409	14	CA931312
c 35	57	38.5	413	14	CA931446
c 36	57	38.5	413	14	CA931603
c 37	57	38.5	414	14	CA927769
c 38	57	38.5	414	14	CA928453
c 39	57	38.5	414	14	CA929036
40	57	38.5	415	13	BU819016
41	57	38.5	416	9	AI164854
42	57	38.5	417	14	CA928113
c 43	57	38.5	417	14	CA929039
c 44	57	38.5	427	14	CA931564
c 45	57	38.5	427	14	CA931687

ALIGNMENTS

RESULT 1  
CB597827/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

CB597827 747 bp mRNA linear EST 03-APR-2003  
AGENCOURT\_12972312 NIH\_MGC\_178 Mus musculus CDNA clone  
IMAGE:30297156 5', mRNA sequence.

CB597827  
CB597827.1 GI:29515683  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 747)

Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasuniishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 KRKEN Mouse ESTs (Konno, H., et al.)

RIKEN Mouse ESTs (Konno, H., et al.)  
Unpublished  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
Five Institute of Physical and Chemical Research (RIKEN)  
1-1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Tel.: 81-43-503-5222  
 Fax: 81-43-503-5216  
 E-mail: genome-res@sc.riken.go.jp,  
 URL: <http://genome.sgc.riken.go.jp/>  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki,  
 N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermotranscription and thermoactivation of the thermolabile enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
 Y. and Hayashizaki, Y. Automated high-throughput plasmid preparation  
 system. *Genome Res.* 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp/>) for  
 further details.

```
/organism= "Mus musculus"
/mol type= "mRNA"
```

```

/motor_type= BLANK
/strain="C57BL/6J"

```

```
/db_xref="taxon:10090"
```

```
/clone="9330145E04"
```

```
/sex="male"
```

```
./tissue_type="diencephalon"
```

```
/dev_stage="adult"
```

```
/lab_host="DH10B"
```

```
/clone_lib="RIKEN full-length enriched, adult male
```

diencephalon"

```
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
```

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken  
Genetic Science Center and Genome Science Laboratory in

Genomic Sciences Center and Genome Science Laboratory in  
 Division of Experimental Animal Research in Riken

RIKEN. DIVISION OF EXPERIMENTAL ANIMAL RESEARCH IN RIKEN  
contributed to prepare mouse tissues 1st strand cDNA was

contributed to prepare mouse tissues. 1st strain CDNA was primed with a primer [5,

primed with a primer [5'-GAGAGAGAAGGATCCAAGAGCTCTTTT TTTT TTTVN 3']. cDNA was

GAGAGAGAGAGGATCCAGAGACCTTTTCTTTTTTAA-3', cDNA was prepared by using trehalose thermo-activated reverse

prepared by using thermase enzyme activated by transcriptase and subsequently enriched for full-length by

transcriptase and subsequently entered for the synthesis of cap-trapper. cDNA went through one round of normalization

to Rot = 10.0 and subtraction to Rot = 185.0. Second

strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGATTCTCGAGTTAATTAAATTATCCCCCCCCCCCC

3' J. cDNA was cloned into the XhoI and BamHI sites.

Vector: a modified pBluescript KS(+) after bulk excision

from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:

**ВамHI "**

47 a	77 c	48 g	66 t
------	------	------	------

ores: 407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187

48.7 238  
Length: 9  
Weight: 9

57.00  
Matches:  
Concurrent (100)

Identity:	75.00%	Conservative:	6
Mismatch:	45.00%	Mismatches:	5

Similarity:	45.00%	Mismatches:	3
	38 51%	Indels:	0

38.51%	INDEXES:	0
10	Gaps:	0

OT  
app:



```

Email: riker@biochem.kth.se.
FEATURES
  source
    Location/Qualifiers
      1..351
        /organism="Populus tremula x Populus tremuloides"
        /mol_type="mRNA"
        /strain="T89"
        /db_xref="taxon:47664"
        /tissue_type="leaf"
        /clone_lib="Populus strain T89 leaves"
BASE COUNT 102 a 55 c 95 g 97 t 2 others
ORIGIN
Alignment Scores:
  Pred. No.: 78.3 Length: 351
  Score: 57.00 Matches: 10
  Percent Similarity: 66.67% Conservative: 6
  Best Local Similarity: 41.67% Mismatches: 8
  Query Match: 38.51% Indels: 0
  DB: 12 Gaps: 0
US-09-684-215A-17 (1-30) x BI070561 (1-351)
QY 1 ThrAlaAspSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 32 ACATCTGGAGCAGAGGATTCGAGTGTTCCTCAAGGAGGATCGAGGTTCTTCGCA 91
QY 21 GlyGlnAlaMet 24
Db 92 GGAGAAATGATG 103
RESULT 6
BI055877
LOCUS CM3-GN0330-120201-735-e08 GN0330 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI055877
VERSION BI055877.1 GI:14463407
KEYWORDS EST.
SOURCE Homo sapiens
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 366)
    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
    Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
    Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
    Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
    ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
    Simpson,A.J.
  Shotgun sequencing of the human transcriptome with ORF expressed
  sequence tags
  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  20202663
  10737800
  Contact: Simpson A.J.G.
  Laboratory of Cancer Genetics
  Ludwig Institute for Cancer Research
  Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
  Brazil
  Tel: +55-11-2704922
  Fax: +55-11-2707001
  Email: asimpson@ludwig.org.br
  This sequence was derived from the FAPESP/LICR Human Cancer Genome
  Project. This entry can be seen in the following URL
  (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&lt2=CM3-GN0330-
  120201-735-e08&t3=2001-02-12&t4=1)
  Seq primer: puc 18 forward
  High quality sequence stop: 366.
  Location/Qualifiers
    1..366
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
FEATURES
  source
    Location/Qualifiers
      1..366
        /organism="Populus tremuloides"
        /mol_type="mRNA"
        /db_xref="taxon:3693"
        /clone_lib="Aspen apex cDNA Library"
        /note="organ: apex"
BASE COUNT 70 a 111 c 64 g 125 t
ORIGIN
Alignment Scores:
  Pred. No.: 83.5 Length: 370
  Score: 57.00 Matches: 10
  Percent Similarity: 66.67% Conservative: 6
  Best Local Similarity: 41.67% Mismatches: 8
  Query Match: 38.51% Indels: 0
  DB: 14 Gaps: 0

```

```

/clone_lib="GN0330"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 82 a 91 c 143 g 50 t
ORIGIN
Alignment Scores:
  Pred. No.: 82.4 Length: 366
  Score: 57.00 Matches: 10
  Percent Similarity: 71.43% Conservative: 5
  Best Local Similarity: 47.62% Mismatches: 6
  Query Match: 38.51% Indels: 0
  DB: 12 Gaps: 0
US-09-684-215A-17 (1-30) x BI055877 (1-366)
QY 3 AlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIleGlyGln 22
Db 184 GCACCGGAGGAGGACAGCTCGATCGCCCTGGCAAGGATTCGTGCTGTGGCGGAG 243
QY 23 Ala 23
Db 244 GCC 246
RESULT 7
CA931030
LOCUS MTU2TA.P1.A09 Aspen apex cDNA Library Populus tremuloides cDNA,
DEFINITION
ACCESSION CA931030
VERSION CA931030.1 GI:27419510
KEYWORDS EST.
SOURCE Populus tremuloides (quaking aspen)
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
    ; eurosids I; Malpighiales; Salicaceae; Populus.
  1 (bases 1 to 370)
    Ranjan,P., Kuo,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
    ,C.-J.
  Expressed sequence tags from Aspen
  Unpublished
  Contact: Tsai C.-J
  Plant Biotech Research Center
  Michigan Technological University, School of Forest Resources &
  Environmental Science
  1400 Townsend Drive, Houghton, MI 49931-1295, USA
  Tel: 906 487 2914
  Fax: 906 487 2915
  Email: chtsa@mtu.edu.
  Location/Qualifiers
    1..370
      /organism="Populus tremuloides"
      /mol_type="mRNA"
      /db_xref="taxon:3693"
      /clone_lib="Aspen apex cDNA Library"
      /note="organ: apex"
BASE COUNT 70 a 111 c 64 g 125 t
ORIGIN
Alignment Scores:
  Pred. No.: 83.5 Length: 370
  Score: 57.00 Matches: 10
  Percent Similarity: 66.67% Conservative: 6
  Best Local Similarity: 41.67% Mismatches: 8
  Query Match: 38.51% Indels: 0
  DB: 14 Gaps: 0

```

KEYWORDS EST.  
SOURCE Populus tremuloides (quaking aspen)

**Spermatophyta: Magnoliophyta; eudicotyledons; core eudicots; rosids**



; eucosids I; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 374)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai ,C-J.

TITLE Expressed sequence tags from Aspen

JOURNAL Unpublished  
COMMENT Contact: Tsai C-J  
Plant Biotech Research Center  
Michigan Technological University, School of Forest Resources &  
Environmental Science  
1400 Townsend Drive, Houghton, MI 49931-1295, USA  
Tel: 906 487 2914  
Fax: 906 487 2915  
Email: chtsal@mtu.edu.

FEATURES source

Location/Qualifiers  
1. .374  
/organism="Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:3693"  
/clone\_lib="Aspen root cDNA Library"  
/notes="Organ: root"

BASE COUNT 69 a 108 c 66 g 131 t  
ORIGIN

Alignment Scores:  
Pred. No.: 84.6 Length: 374  
Score: 57.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 8  
Query Match: 38.51% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x CA927770 (1-374)

Qy 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 146 ACATCTGGAGCAGGAGTTCGAAGTGTTCACAGGAGGATCAGGGTTCGCATTCTTCGCA 87

Qy 21 GlyGlnAlaMet 24

Db 86 GGAGAAATGATG 75

RESULT 14

CA928449/c

LOCUS MTU6TR.P2.G12 Aspen root cDNA Library Populus tremuloides cDNA, 374 bp mRNA linear EST 30-DEC-2002

DEFINITION mRNA sequence.

ACCESSION CA928449

VERSION CA928449.1 GI:27416928

KEYWORDS EST.

SOURCE Populus tremuloides (quaking aspen)

ORGANISM Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

REFERENCE 1 (bases 1 to 374)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai ,C-J.

TITLE Expressed sequence tags from Aspen

JOURNAL Unpublished  
COMMENT Contact: Tsai C-J  
Plant Biotech Research Center  
Michigan Technological University, School of Forest Resources &  
Environmental Science  
1400 Townsend Drive, Houghton, MI 49931-1295, USA  
Tel: 906 487 2914  
Fax: 906 487 2915  
Email: chtsal@mtu.edu.

FEATURES source

Location/Qualifiers  
1. .374  
/organism="Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:3693"

/clone\_lib="Aspen root cDNA Library"  
/notes="Organ: root"

BASE COUNT 71 a 110 c 64 g 129 t  
ORIGIN

Alignment Scores:

Pred. No.: 84.6 Length: 374  
Score: 57.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 8  
Query Match: 38.51% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x CA928449 (1-374)

Qy 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 146 ACATCTGGAGCAGGAGTTCGAAGTGTTCACAGGAGGATCAGGGTTCGCATTCTTCGCA 87

Qy 21 GlyGlnAlaMet 24

Db 86 GGGGAATGATG 75

RESULT 15

CA928493/c

LOCUS MTU6TR.P3.C12 Aspen root cDNA Library Populus tremuloides cDNA, 374 bp mRNA linear EST 30-DEC-2002

DEFINITION mRNA sequence.

ACCESSION CA928493

VERSION CA928493.1 GI:27416972

KEYWORDS EST.

SOURCE Populus tremuloides (quaking aspen)

ORGANISM Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

REFERENCE 1 (bases 1 to 374)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai ,C-J.

TITLE Expressed sequence tags from Aspen

JOURNAL Unpublished

COMMENT Contact: Tsai C-J

Plant Biotech Research Center

Michigan Technological University, School of Forest Resources &

Environmental Science

1400 Townsend Drive, Houghton, MI 49931-1295, USA

Tel: 906 487 2914

Fax: 906 487 2915

Email: chtsal@mtu.edu.

FEATURES source

Location/Qualifiers

1. .374

/organism="Populus tremuloides"

/mol\_type="mRNA"

/db\_xref="taxon:3693"

/clone\_lib="Aspen root cDNA Library"

/notes="Organ: root"

BASE COUNT 70 a 110 c 65 g 129 t

ORIGIN

Alignment Scores:

Pred. No.: 84.6 Length: 374  
Score: 57.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 8  
Query Match: 38.51% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x CA928493 (1-374)

Qy 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 146 ACATCTGGAGCAGGAGTTCGAAGTGTTCACAGGAGGATCAGGGTTCGCATTCTTCGCA 87

Qy 21 GlyGlnAlaMet 24  
   ||||:|  
Db 86 GGAGAAATGATG 75

Search completed: September 5, 2003, 11:54:36  
Job time : 520.276 secs